OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: August 20, 2004, 03:13:58 ; Search time 8782.76 Seconds
(without alignments)
10940.929 Million cell updates/sec

Title: US-09-894-749-1
Perfect score: 2217
Sequence: 1 gaattcggcttccatcctaa........aaaaaaaaaggggcggccgc 2217

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 347027 seqs, 2167151699 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pri:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: gb_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2217	100.0	2217	6	AR165090	AR165090 Sequence
2	2217	100.0	2217	6	BD265598	BD265598 RGS-conta
3	2217	100.0	2217	6	AR214935	AR214935 Sequence
4	2068	93.3	2138	9	AF076642	AF076642 Homo sapi
5	2064.4	93.1	2136	9	AK057114	AK057114 Homo sapi
6	2063.2	93.1	2164	9	BC020632	BC020632 Homo sapi
7	2041.6	92.1	2144	6	AX299976	AX299976 Sequence
8	2041.6	92.1	2144	6	AF268036	AF268036 Homo sapi
9	1753.2	79.1	1840	6	AX299975	AX299975 Sequence
10	1581	71.3	94175	9	AL513175	AL513175 Human DNA
11	1581	71.3	107919	9	AL391274	AL391274 Homo sapi
12	1448.2	65.3	1486	6	AX299963	AX299963 Sequence
13	596	26.9	1164	6	AR165091	AR165091 Sequence
14	596	26.9	1164	6	BD265699	BD265699 RGS-conta
15	596	26.9	1164	6	AR214936	AR214936 Sequence
16	596	26.9	1937	10	AB042807	AB042807 Mus muscu
17	585.6	26.4	1399	10	AF302685	AF302685 Mus muscu
18	441.6	19.9	448	11	G52997	G52997 SHGC-81782
19	252.4	11.3	110000	2	AC112625_1	Continuation (2 of
20	227.4	10.3	218670	10	AC102163_3	AC102163 Mus muscu
21	227.2	10.2	241	6	AX299968	AX299968 Sequence
22	218	9.8	154968	9	AL596342	AL596342 Human DNA
23	160.4	7.2	494	4	AF472577	AF472577 Ovis arie
24	158.4	7.1	408	6	AX710064	AX710064 Sequence
25	155.6	7.0	636	9	AF493926	AF493926 Homo sapi
26	155.6	7.0	636	9	BT007065	BT007065 Homo sapi
27	155.6	7.0	636	12	BT008081	BT008081 Synthetic
28	155.6	7.0	1345	6	AR270715	AR270715 Sequence
29	155.6	7.0	1345	6	AR380698	AR380698 Sequence
30	155.6	7.0	1345	6	AX663626	AX663626 Sequence
31	155.6	7.0	1345	9	HUMGOS8PPC	L13463 Human helix
32	155.6	7.0	1377	9	BC007049	BC007049 Homo sapi
33	153.2	6.9	546	9	AF493929	AF493929 Homo sapi
34	153.2	6.9	744	6	AR075100	AR075100 Sequence
35	153.2	6.9	744	6	BD091623	BD091623 Human reg
36	153.2	6.9	1285	9	AF030108	AF030108 Homo sapi
37	153.2	6.9	1691	9	AF159570	AF159570 Homo sapi
38	153.2	6.9	1710	9	BC030059	BC030059 Homo sapi
39	153.2	6.9	1923	6	BC074349	BC074349 50 human
40	153.2	6.9	2076	9	AB008109	AB008109 Homo sapi
41	153.2	6.9	2190	6	AR380250	AR380250 Sequence
42	153.2	6.9	5683	9	HSM805729	BX537427 Homo sapi
43	152.4	6.9	685	10	AF321837	AF321837 Rattus no
44	152.4	6.9	1363	10	AY043246	AY043246 Rattus no
45	152.4	6.9	1458	10	BC061969	BC061969 Rattus no

ALIGNMENTS

RESULT 1
AR165090
LOCUS   AR165090   2217 bp   DNA   linear   PAT 17-OCT-2001
DEFINITION   Sequence 1 from patent US 6274362.
ACCESSION   AR165090
VERSION   AR165090.1 GI:16238486
KEYWORDS   .
SOURCE   Unknown.
ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 2217)
AUTHORS   Hodge,M.R. and Yowe,D.
TITLE   RGS-containing molecules and uses thereof
JOURNAL   Patent: US 6274362-A 1 14-AUG-2001;
FEATURES   Location/Qualifiers

Pred. No. is the number of results predicted by chance to have a

```
source      1..2217
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN

Query Match    100.0%;  Score 2217;  DB 6;  Length 2217;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 2217;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy     1 GAATTCGGCTTCCATCCTAATACGACTCACTATAGGGCGTCGAGCGCCCGGGCAGGT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GAATTCGGCTTCCATCCTAATACGACTCACTATAGGGCGTCGAGCGCCCGGGCAGGT 60

Qy    61 ATAACTTTTATTCTACTACATGTATATGTGAATACTATTAATAATAAATGAACTAGGGAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ATAACTTTTATTCTACTACATGTATATGTGAATACTATTAATAATAAATGAACTAGGGAAG 120

Qy   121 GATGTAATAAATTAGACATCTCTTCATTTGAATCAAAGAGAAAAAACTTTTTCAAGTAATACATGGT 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GATGTAATAAATTAGACATCTCTTCATTTGAATCAAAGAGAAAAAACTTTTTCAAGTAATACATGGT 180

Qy   181 TTTTCAAATAAATAAATGTGAATCAAAGAGAAAAATCAGAGCTAAGAAAAAAGAAAATAGA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TTTTCAAATAAATAAATGTGAATCAAAGAGAAAAATCAGAGCTAAGAAAAAAGAAAATAGA 240

Qy   241 TCAGGAAAAGAAGAAAGAAACAAGCAAAGAAGCCAAATCAGAGCTAAGAAAAAAGAAAATAGA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TCAGGAAAAGAAGAAAGAAACAAGCAAAGAAGCCAAATCAGAGCTAAGAAAAAAGAAAATAGA 300

Qy   301 CTAAGTCTTCTTGTGCAGAAAACCTGAGTTCATGAAGACACCCGCTCCAGTGATCTCGGG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTAAGTCTTCTTGTGCAGAAAACCTGAGTTCATGAAGACACCCGCTCCAGTGATCTCGGG 360

Qy   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT 420

Qy   421 GACAAACTGCTTCCCATAGAGATGGACTAGAGCTTTTACCAGATTTCTTAAAACTGAA 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GACAAACTGCTTCCCATAGAGATGGACTAGAGCTTTTACCAGATTTCTTAAAACTGAA 480

Qy   481 TTCAGTGAAGAAATATTGAATTTGAATTTTGGATAGCCTGTGTGAAGATTTCAAGAAAAGCAAGGGA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TTCAGTGAAGAAATATTGAATTTGAATTTTGGATAGCCTGTGTGAAGATTTCAAGAAAAGCAAGGGA 540

Qy   541 CCTCAACAAATTCACCTTAAAGCAAAGCAATATGAGAAATTTATACAGAAACAGCACTCAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCTCAACAAATTCACCTTAAAGCAAAGCAATATGAGAAATTTATACAGAAACAGCACTCAA 600

Qy   601 CCAAAAGAGGTTAACCTTGATTTCACACAAAGAAGAGTCAATTACACAAAGAGTCATGATGCC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CCAAAAGAGGTTAACCTTGATTTCACACAAAGAAGAGTCAATTACACAAAGAGTCATGATGCC 660

Qy   661 CCTACCCTCCACAGTTTTGATGCGTCGACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CCTACCCTCCACAGTTTTGATGCGTCGACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC 720

Qy   721 AGTTATACACGTTTCTGAAATCTGACACTCATTAGACTTAGACTTATTTGCTCATTAGATAATTC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AGTTATACACGTTTCTGAAATCTGACACTCATTAGACTTAGACTTATTTGCTCATTAGATAATTC 780

Qy   781 AGACCAACAAATCTTAGGAGACAGCATCACGCTCATTTACCTGCCAATGAATTCCAAGATGTA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGACCAACAAATCTTAGGAGACAGCATCACGCTCATTTACCTGCCAATGAATTCCAAGATGTA 840

Qy   841 CAATCAGATGTTGCCATTGGTTATAAAGAAAAATTGATTTTGCTCATTTTATGACAAAC 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CAATCAGATGTTGCCATTGGTTATAAAGAAAAATTGATTTTGCTCATTTTATGACAAAC 900

Qy   901 TTTATACATCTGCTTCTAACATATCGCATGTTAATGTTAAGATTGGTCCCATCCTTTAAA 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTTATACATCTGCTTCTAACATATCGCATGTTAATGTTAAGATTGGTCCCATCCTTTAAA 960
```

```
Qy   961 CTGAAATATGTCATGTGAAATTATTTAAAAATGTAAAAAACAAAACTTTCTGCTAACAAA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CTGAAATATGTCATGTGAAATTATTTAAAAATGTAAAAAACAAAACTTTCTGCTAACAAA 1020

Qy  1021 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCATTTGATGTCATTCCATTTA 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCATTTGATGTCATTCCATTTA 1080

Qy  1081 TAATCAGAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAAGTAATAATGTTT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TAATCAGAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAAGTAATAATGTTT 1140

Qy  1141 TATAAGATTGTAGAGTTAAGTAAAAGTTAAGCTTTTGCAAAAGTTGTCAAAAGTTCAAACA 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TATAAGATTGTAGAGTTAAGTAAAAGTTAAGCTTTTGCAAAAGTTGTCAAAAGTTCAAACA 1200

Qy  1201 AAAGTCTAGTGGGATTTTTTACCAAAGCAGCATAAATATGTGTCATAATGAAATGTTCTCTTTTTTTGGT 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 AAAGTCTAGTGGGATTTTTTACCAAAGCAGCATAAATATGTGTCATAATGAAATGTTCTCTTTTTTTGGT 1260

Qy  1261 ACTCAGATATCCAAATGTTCAGATAGCATTTTCATAAATGAAATGTTCTTTTTTTTGGT 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 ACTCAGATATCCAAATGTTCAGATAGCATTTTCATAAATGAAATGTTCTTTTTTTTGGT 1320

Qy  1321 AATAGTGTAGAAGTGATCTCGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 AATAGTGTAGAAGTGATCTCGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT 1380

Qy  1381 ACTACTAACCCTGTCCCAGAAATAGTAATATCACCTCTAGTTATAAGCAACAGCAGCAACAGGAA 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 ACTACTAACCCTGTCCCAGAAATAGTAATATCACCTCTAGTTATAAGCAACAGCAGCAACAGGAA 1440

Qy  1441 CTTTTGTGAAGACACATTCCACATTGAACTCATTCCTAAGTGAACATGGACTACCCAGTTAT 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 CTTTTGTGAAGACACATTCCACATTGAACTCATTCCTAAGTGAACATGGACTACCCAGTTAT 1500

Qy  1501 ACTGAGAATAAGATCCACATTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGTAGGGA 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 ACTGAGAATAAGATCCACATTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGTAGGGA 1560

Qy  1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGTAGGGA 1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGTAGGGA 1620

Qy  1621 ACTAAGTAGACTACCTTATCCGGCTAAGAAAACTGCTACTAAACTATTAGGCCATCA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 ACTAAGTAGACTACCTTATCCGGCTAAGAAAACTGCTACTAAACTATTAGGCCATCA 1680

Qy  1681 ATGGCTTGAATAAAACCAGAGAGGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 ATGGCTTGAATAAAACCAGAGAGGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG 1740

Qy  1741 AATTGGGGTAGAAATCAGAAATGAGGGGAAGAAGCAAGGAGTCTAAGGCCCTAGC 1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 AATTGGGGTAGAAATCAGAAATGAGGGGAAGAAGCAAGGAGTCTAAGGCCCTAGC 1800

Qy  1801 GATTTGGGCATCGCCACATTGGTTATCTCATTGATTAATTCTAT 1860
         |||||||||||||||||||||||||||||||||||||||||||
Db  1801 GATTTGGGCATCGCCACATTGGTTATCTCATTGATTAATTCTAT 1860

Qy  1861 TTGTTAAGCAAATCTCCTTAAGTAATTATTATTCAAATAAGATTATACTATACATCTAT 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 TTGTTAAGCAAATCTCCTTAAGTAATTATTATTCAAATAAGATTATACTATACATCTAT 1920

Qy  1921 ATGTCACTGTTTAAAAATGCCATTGGTTATAAAGATATTTAATTTTAATGTGTTACATGGCTCTGTAAATATT 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 ATGTCACTGTTTAAAAATGCCATTGGTTATAAAGATATTTAATTTTAATGTGTTACATGGCTCTGTAAATATT 1980

Qy  1981 TGTATTTAAAAATGCCATGCATAGGCTTGGAAATTTAATGTTAGTTGAAATGTAAAAT 2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 TGTATTTAAAAATGCCATGCATAGGCTTGGAAATTTAATGTTAGTTGAAATGTAAAAT 2040

Qy  2041 GTGAAAACTTTAGTACATTTGTAGTAATAAATATTTTAACTTCATTCATTCACAGTTAAGT 2100
```

```
Db   2041 GTGAAAACTTAGATCATTTGTAGTAATAAATATTTTAACTTCATTCATTACACAGTTAAGT 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2101 TTATCTGCAATAAAAGCTCGACTGACTGATTATCCTTCCTATTATATGTAATAAGGA 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 TTATCTGCAATAAAAGCTCGACTGACTGATTATCCTTCCTATTATATGTAATAAGGA 2160

Qy   2161 ATAAACATTTTCTTCTTTTTAGAGTAAAAAAAAAAAAAAAAAAAGGGGCGGCCGC 2217
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ATAAACATTTTCTTCTTTTTAGAGTAAAAAAAAAAAAAAAAAAAGGGGCGGCCGC 2217
```

RESULT 2
BD265698
LOCUS   BD265698   2217 bp   DNA   linear   PAT 17-JUL-2003
DEFINITION   RGS-containing molecules and uses thereof.
ACCESSION   BD265698
VERSION   BD265698.1 GI:33075466
KEYWORDS   JP 2002535979-A/1.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2217)
AUTHORS   Hodge.M.R. and Yowe.D.
TITLE   RGS-containing molecules and uses thereof
JOURNAL   Patent: JP 2002535979-A 1 29-OCT-2002;
             MILLENNIUM PHARMACEUTICALS INC
COMMENT   OS   Homo sapiens (human)
             PN   JP 2002535979-A/1
             PD   29-OCT-2002
             PF   04-FEB-2000 JP 2000597306
             PF   04-FEB-1999 US   09/244314
             PI   MARTIN R HODGE,DAVID YOWE
             PC   C12N15/09,C07K14/47,C07K16/18,C12N5/10,C12P21/02,C12Q1/02, PC
             C12Q1/68,
             PC   G01N33/53,G01N33/566,G01N33/53,G01N15/00,C12N5/00 CC
             RGS-containing molecules and uses thereof
FEATURES   FH   Key   Location/Qualifiers
             FT   CDS   (160)..(867).
             source   1..2217
                             /organism="Homo sapiens"
                             /mol_type="genomic DNA"
                             /db_xref="taxon:9606"
ORIGIN

Query Match   100.0%; Score 2217; DB 6; Length 2217;
Best Local Similarity 100.0%; Pred. No.0;
Matches 2217; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 GAATTCGGCTTCCATCCAATACGACTCACTATAGGGCTCGAGCGGCCGCCCCCGGGCAGGT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAATTCGGCTTCCATCCAATACGACTCACTATAGGGCTCGAGCGGCCGCCCCCGGGCAGGT   60

Qy     61 ATAACTTTTTATTCTACTATGTATGTATATGATGGAATAGTATTAATAAATGAACTAGGGAAG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATAACTTTTTATTCTACTATGTATGTATATGATGGAATAGTATTAATAAATGAACTAGGGAAG  120

Qy    121 GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAGATGGAAACAACATTGCTTTTC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAGATGGAAACAACATTGCTTTTC  180

Qy    181 TTTTTTCAAATAATAATGTGAATCAAAGAGAAAAAAACTTTTCAAGTAATACATGGT  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTTTTTCAAATAATAATGTGAATCAAAGAGAAAAAAACTTTTCAAGTAATACATGGT  240

Qy    241 TCAGGAAAAGAAGAAACAACAAGCAACGCAAATCAGAGCCAAATCAGAGAAAAATGAA  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCAGGAAAAGAAGAAACAACAAGCAACGCAAATCAGAGCCAAATCAGAGAAAAATGAA  300

Qy    301 CTAAGTCTTCTGTGCAGAAACCTGAGTTCATGAAGACACCCCGCTCCAGTAGATCTGGG  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTAAGTCTTCTGTGCAGAAACCTGAGTTCATGAAGACACCCCGCTCCAGTAGATCTGGG  360

Qy    361 CACTTGGCCAAAGAACAAGAGCTCCCCTGAAGAGGCAGTCAGTGAAATGGGTGAATCATT  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACTTGGCCAAAGAACAAGAGCTCCCCTGAAGAGGCAGTCAGTGAAATGGGTGAATCATT  420

Qy    421 GACAAACTGCTTTCCATAGAGATGGACTAGAGGCCTTTACCAGATTCTTAAAACTGAA  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GACAAACTGCTTTCCATAGAGATGGACTAGAGGCCTTTACCAGATTCTTAAAACTGAA  480

Qy    481 TTCAGTGAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTCAGTGAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA  540

Qy    541 CCTCAACAAATTCACCTTAAACACAAAGCAATATATGACAAATTTATACAGACTGTGCC  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCTCAACAAATTCACCTTAAACACAAAGCAATATATGACAAATTTATACAGACTGATGCC  600

Qy    601 CCAAAAGAGGTTAACCTTGATTTTCACCACAAAGAAGTCATTACAAACAGCATCACTCCAA  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CCAAAAGAGGTTAACCTTGATTTTCACCACAAAGAAGTCATTACAAACAGCATCACTCAA  660

Qy    661 CCTACCCTCCACAGTTTGATGCTGCACAAGCAGAGTGTATCAGCTCATGGAACAGAGAC  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCTACCCTCCACAGTTTGATGCTGCACAAGCAGAGTGTATCAGCTCATGGAACAGAGAC  720

Qy    721 AGTTATACACGTTTCTGAAATCTGACATCTATTAGACTGATGAGGAGGAGACCTCAG  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AGTTATACACGTTTCTGAAATCTGACATCTATTAGACTGATGAGGAGGAGACCTCAG  780

Qy    781 AGACCAACAAATCTTAGGAGACGATCACGCTCATTACCTGCAATGAATTCCAAGATGTA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGACCAACAAATCTTAGGAGACGATCACGCTCATTACCTGCAATGAATTCCAAGATGTA  840

Qy    841 CAATCAGATGTTGCCATTGGTTATAAGAGAAATTGATTTTGCTCATTTTATGACAAAC  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CAATCAGATGTTGCCATTGGTTATAAAGAGAAATTGATTTTGCTCATTTTATGACAAAC  900

Qy    901 TTATACATCTGCTTCTAACATATCGCATGTTATGTTAAGAGATTTGGTCCCATCCTTTAAA  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTATACATCTGCTTCTAACATATCGCATGTTATGTTAAGAGATTTGGTCCCATCCTTTAAA  960

Qy    961 CTGAAATATGTCATGTGAAATTATTTAAAAATGTAAAACAAAACTTTCTGCTAACAAA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CTGAAATATGTCATGTGAAATTATTTAAAAATGTAAAACAAAACTTTCTGCTAACAAA 1020

Qy   1021 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATTGTCATTCATTA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATTGTCATTCATTA 1080

Qy   1081 TAATCAGAAAAAACTTATTTCTTAATCAAAAAGGCAGTACAAAAAAGTAATAATGTTT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TAATCAGAAAAAACTTATTTCTTAATCAAAAAGGCAGTACAAAAAAGTAATAATGTTT 1140

Qy   1141 TATTAAGATTGTAGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTTGTCAAAGTTCAAACA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TATTAAGATTGTAGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTTGTCAAAGTTCAAACA 1200

Qy   1201 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAATGTGTTATATAAAACATAATAAT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAATGTGTTATATAAAACATAATAAT 1260

Qy   1261 ACTCAGATATCCAAATGTTCAGTAGCATTTTTCAAGTAATACATGGT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ACTCAGATATCCAAATGTTCAGTAGCATTTTTCAAGTAATACATGGT 1320

Qy   1321 AATAGTGTAGAGTGATCTGGTTCTTCAATGGGAGATGAAGAACATTTATTATTGGGTT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AATAGTGTAGAGTGATCTGGTTCTTCAATGGGAGATGAAGAACATTTATTATTGGGTT 1380

Qy   1381 ACTACTAACCCTGTCCCAAGAATAGTAATATCCTCAGTATTATAAGCCAGCAACAGGAA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ACTACTAACCCTGTCCCAAGAATAGTAATATCCTCAGTATTATAAGCCAGCAACAGGAA 1440
```

```
Query Match         100.0%;  Score 2217;  DB 6;  Length 2217;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 2217;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy     1 GAATTCGGCTTCCATCCTAATACGACTCACTATAGGGCTCGAGCCGGCCGCCGGGCCAGGT   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GAATTCGGCTTCCATCCTAATACGACTCACTATAGGGCTCGAGCCGGCCGCCGGGCCAGGT   60

Qy    61 ATAACTTTTATTCTACTATGTATATGTATGGGAAGTAGTATTAATAAATGACTAGGGAAG  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ATAACTTTTATTCTACTATGTATATGTATGGGAAGTAGTATTAATAAATGACTAGGGAAG  120

Qy   121 GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGGAAGATGGAAACAACATTGCTTTTC  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGGAAGATGGAAACAACATTGCTTTTC  180

Qy   181 TTTTTCAAATAAATATGTGAATCAAAGAAAAAACTTTTTCAAGTTAATACATGGT  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TTTTTCAAATAAATATGTGAATCAAAGAAAAAACTTTTTCAAGTTAATACATGGT  240

Qy   241 TCAGGAAAGAAGAAACAAGCAAATCAGAGCTAAGGAAATAGA  300
         ||||||||||||||||||||||||||||||||||||||||||
Db   241 TCAGGAAAGAAGAAACAAGCAAATCAGAGCTAAGGAAATAGA  300

Qy   301 CTAAGTCTTCTGTGCGAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGG  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTAAGTCTTCTGTGCGAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGG  360

Qy   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTT  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTT  420

Qy   421 GACAAACTGCTTCCCATAGAGATGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA  480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GACAAACTGCTTCCCATAGAGATGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA  480

Qy   481 TTCAGTGAAGAAATATTGAATTTGGATAGCCTGTGAAGATTCAAGAAAGAAGCAAGGGA  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TTCAGTGAAGAAATATTGAATTTGGATAGCCTGTGAAGATTCAAGAAAGAAGCAAGGGA  540

Qy   541 CCTCAACAAATCACCTTAAAGCAAAAGCAATATATGAAGCAATATGAGAAATTTATACAGACTGATGCC  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCTCAACAAATCACCTTAAAGCAAAAGCAATATATGAAGCAATATGAGAAATTTATACAGACTGATGCC  600

Qy   601 CCAAAAGAGGTAACCTTGATTTTCACACAAAAGAAGTCATTACAAAACAGCATCACTCAA  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CCAAAAGAGGTAACCTTGATTTTCACACAAAAGAAGTCATTACAAAACAGCATCACTCAA  660

Qy   661 CCTACCCTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGTCATGGAACAAGAC  720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CCTACCCTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGTCATGGAACAAGAC  720

Qy   721 AGTTATACACGTTTCTGAAATCTGACATCTATTTGACTTGATGGAAGGAAGGAGACCTCAG  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AGTTATACACGTTTCTGAAATCTGACATCTATTTGACTTGATGGAAGGAAGGAGACCTCAG  780

Qy   781 AGACCAACAAATCTTAGGAGACGATCGTCATTACCTGCAATGAATTCCAAGATGTA  840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGACCAACAAATCTTAGGAGACGATCGTCATTACCTGCAATGAATTCCAAGATGTA  840

Qy   841 CAATCAGATGTTGCCATTGGTTATAAGAAATTGATTTTGCTCATTTTATGACAAAC  900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CAATCAGATGTTGCCATTGGTTATAAGAAATTGATTTTGCTCATTTTATGACAAAC  900

Qy   901 TTATACATCTGCTTCTAACATATCCGATGTTTATGTTAAGATTTGGTCCCATCCTTTAAA  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTATACATCTGCTTCTAACATATCCGATGTTTATGTTAAGATTTGGTCCCATCCTTTAAA  960

Qy   961 CTGAAATATGTCATGTGAAATTATTTAAAAATGTAAAAACAAAACTTTCTGCTAACAAA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CTGAAATATGTCATGTGAAATTATTTAAAAATGTAAAAACAAAACTTTCTGCTAACAAA 1020

Qy  1441 CTTTTGTGAAGACACATTCATCTCTACAGAACTTCAGATTAAATAATATCTAGATTAATG 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 CTTTTGTGAAGACACATTCATCTCTACAGAACTTCAGATTAAATAATATCTAGATTAATG 1500

Qy  1501 ACTGAGAATAAGATCCACATTGAACTCATTCCTAAGTGAACATGGAACGTACCCAGTTAT 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 ACTGAGAATAAGATCCACATTGAACTCATTCCTAAGTGAACATGGAACGTACCCAGTTAT 1560

Qy  1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGTAGGGA 1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGTAGGGA 1620

Qy  1621 ACTAAGTAGACTACCCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGCCATCA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 ACTAAGTAGACTACCCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGCCATCA 1680

Qy  1681 ATGGCTTGAATAAAAACCAGAGAAGGTTTTCCCAGGACGTCTCATGTTGGCCCTTTAG 1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 ATGGCTTGAATAAAAACCAGAGAAGGTTTTCCCAGGACGTCTCATGTTGGCCCTTTAG 1740

Qy  1741 AATTGGGTAGAAATCAGAAATGAGATGAGGGGAAGAGCAAGCAAGGAGTCTAAGGCCCTAGC 1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 AATTGGGTAGAAATCAGAAATGAGATGAGGGGAAGAGCAAGCAAGGAGTCTAAGGCCCTAGC 1800

Qy  1801 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGATTATATTC 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGATTATATTC 1860

Qy  1861 TTGTTAAGCAAATCTCCTTAAGAGATAATTATTATTATTCAAATAAGATTATACATCTAT 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 TTGTTAAGCAAATCTCCTTAAGAGATAATTATTATTATTCAAATAAGATTATACATCTAT 1920

Qy  1921 ATGTCACTGTTTTAAAAGAGATATTTAATTTTAATGTGTTACATGGTCTGTAAATATT 1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 ATGTCACTGTTTTAAAAGAGATATTTAATTTTAATGTGTTACATGGTCTGTAAATATT 1980

Qy  1981 TGTATTTAAAAAATGCCATGCATTAGGCTTTGGAAATTTAATGTTAGTTGAAATGTAAAAT 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 TGTATTTAAAAAATGCCATGCATTAGGCTTTGGAAATTTAATGTTAGTTGAAATGTAAAAT 2040

Qy  2041 GTGAAAACTTTAGATCATTGTAGTAATAAATATTTTAACTTCATTCATACAGTTAAGT 2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 GTGAAAACTTTAGATCATTGTAGTAATAAATATTTTAACTTCATTCATACAGTTAAGT 2100

Qy  2101 TTATCTGACAATAAAGCTCTGACTGAATGTTGATTATCCTTCCTATTATGTAATAAGGA 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 TTATCTGACAATAAAGCTCTGACTGAATGTTGATTATCCTTCCTATTATGTAATAAGGA 2160

Qy  2161 ATAAACATTTTCTTCTTTTTAGAGTAAAAAAAAAAAAAAAAAAAAAAAGGGCCGCCGC 2217
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 ATAAACATTTTCTTCTTTTTAGAGTAAAAAAAAAAAAAAAAAAAAAAAGGGCCGCCGC 2217
```

```
RESULT 3
AR214935
LOCUS       AR214935                2217 bp    DNA     linear      PAT 25-SEP-2002
DEFINITION  Sequence 1 from patent US 6410240.
ACCESSION   AR214935
VERSION     AR214935.1  GI:23312888
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2217)
  AUTHORS   Hodge,M.R. and Yowe,D.
  TITLE     RGS-containing molecules and uses thereof
  JOURNAL   Patent: US 6410240-A 1 25-JUN-2002;
FEATURES             Location/Qualifiers
     source          1..2217
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

```
Db   2101 TTATCTGACAATAAAAGCTCTGACTGAATGTTCGATTATCCTTCCTATTATGTAATAAGGA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   2161 ATAAACATTTCTCTTTAGAGTAAAAAAAAAAAAAAAAAAAAAAAAGGGGCCGC 2217
          ||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ATAAACATTTCTCTTTAGAGTAAAAAAAAAAAAAAAAAAAAAAAAGGGGCCGC 2217
```

RESULT 4
AF076642
LOCUS   AF076642   2138 bp   mRNA   linear   PRI 01-JUL-2000
DEFINITION Homo sapiens regulator of G-protein signaling 13 mRNA, complete
            cds.
ACCESSION   AF076642
VERSION   AF076642.1 GI:8885627
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2138)
  AUTHORS   Zhang,W., Wan,T., Yuan,Z., He,L. and Cao,X.
  TITLE   A novel regulator of G-protein signaling
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 2138)
  AUTHORS   Zhang,W., Wan,T., Yuan,Z., He,L. and Cao,X.
  TITLE   Direct Submission
  JOURNAL   Submitted (07-JUL-1998) Department of Immunology, Shanghai
            Brilliance Biotechnology Institute & Second Military Medical
            University, 800 Xiangyin Road, Shanghai 200433, P.R. China
FEATURES   Location/Qualifiers
     source   1..2138
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="SBBI05"
     CDS   151..858
                     /note="RGS13"
                     /codon_start=1
                     /product="regulator of G-protein signaling 13"
                     /protein_id="AAF80227.1"
                     /db_xref="GI:8885628"
                     /translation="METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEK
                     RNRLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTR
                     FLKTEFSEENIEFWIACEDFKKSKGPQQIHIKAKAIYEKFIQTDAPKEVNDFHTKEV
                     ITNSITQPTLHSFDAAQSRVYLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRS
                     FTCNEFQDVQSDVAIWL"
ORIGIN

Query Match   93.3%; Score 2068; DB 9; Length 2138;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 2068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   61 ATAACTTTTATTCTACTACTGTATATGTATGGAATGATTATTAATAAATGAACTAGGGAAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   52 ATAACTTTTATTCTACTACTGTATATGTATGGAATGATTATTAATAAATGAACTAGGGAAG 111

Qy   121 GATGTAATAAATTAGACATCTCTTCATTTTAGAGAAGAAGATGGAAGAACAACATTGCTTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112 GATGTAATAAATTAGACATCTCTTCATTTTAGAGAAGAAGATGGAAGAACAACATTGCTTTTC 171

Qy   181 TTTTCTCAAATAATAATGTGTGAATCAAAGAAAAAAACTTTTTCAAGTTAATAACATGGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   172 TTTTCTCAAATAATAATATGTGTGAATCAAAGAAAAAAACTTTTTCAAGTTAATAACATGGT 231

Qy   241 TCAGGAAGAGAGAAACAAGCAAGACCAAAATCAGAGCTAAGGAAAAAAGAAATAGA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   232 TCAGGAAGAGAGAAACAAGCAAGACCAAAATCAGAGCTAAGGAAAAAAGAAATAGA 291

Qy   301 CTAAGTCTTCTTGTGCAGAAAACCTGAGTTTCATGAAGACACACCGCCTCCCAGTAGATCTGGG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   292 CTAAGTCTTCTTGTGCAGAAAACCTGAGTTTCATGAAGACACACCGCCTCCCAGTAGATCTGGG 351
```

Qy  1021 ATACATACAGTATCTGCCGCTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTA 1080
Db  1021 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTA 1080

Qy  1081 TAATCAGAAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAGTAATAAATGTTT 1140
Db  1081 TAATCAGAAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAGTAATAAATGTTT 1140

Qy  1141 TATAAGATTGTGAGAGTAAGTAAGTAAAAGTTTTGCAAAGTTGTGTCAAAAGTTCAAACA 1200
Db  1141 TATAAGATTGTGAGAGTAAGTAAGTAAAAGTTTTGCAAAGTTGTGTCAAAAGTTCAAACA 1200

Qy  1201 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAATATGTGTATATAAACATAATAAT 1260
Db  1201 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAATATGTGTATATAAACATAATAAT 1260

Qy  1261 ACTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAAATGTTCTCTTTTTTTTTGGT 1320
Db  1261 ACTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAAATGTTCTCTTTTTTTTTTGT 1320

Qy  1321 AATAGTGTAGAAGTGATCTGGTTCTTCTTACAATGGAGATGAAGAACATTTATTATTGGGTT 1380
Db  1321 AATAGTGTAGAAGTGATCTGGTTCTTCTTACAATGGAGATGAAGAACATTTATTATTGGGTT 1380

Qy  1381 ACTACTAACCCTGTCCCAAGAATAGTAATACACTCTAGTTATAAGCCAGCAACAGGAA 1440
Db  1381 ACTACTAACCCTGTCCCAAGAATAGTAATACACTCTAGTTATAAGCCAGCAACAGGAA 1440

Qy  1441 CTTTTGTGAAGACACATCATCTCTACAGAACTTCAGATTAAATATAATTCTAGATTAATG 1500
Db  1441 CTTTTGTGAAGACACATCATCTCTACAGAACTTCAGATTAAATATAATTCTAGATTAATG 1500

Qy  1501 ACTGAGAATAAGATCCACATTTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT 1560
Db  1501 ACTGAGAATAAGATCCACATTTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT 1560

Qy  1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGGTAGGGA 1620
Db  1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGGTAGGGA 1620

Qy  1621 ACTAAGTAGAGCTACCTTATCACCGGCTAGAGAAAACTTGCTACTACTAAACTATTAGGCCATCA 1680
Db  1621 ACTAAGTAGAGCTACCTTATCACCGGCTAGAGAAAACTTGCTACTACTAAACTATTAGGCCATCA 1680

Qy  1681 ATGGCTTGAATAAAAACCAGAGAGAAGGTTTTTCCCGGAGACGTCTCATGTTTGGCCCTTTAG 1740
Db  1681 ATGGCTTGAATAAAAACCAGAGAGAAGGTTTTTCCCGGAGACGTCTCATGTTTGGCCCTTTAG 1740

Qy  1741 AATTGGGGTAGAAATCAGAAATGAGAATGAGATGAGGGGAAGAAGCAAGGAGTCAAGGCCCTAGC 1800
Db  1741 AATTGGGGTAGAAATCAGAAATGAGAATGAGATGAGGGGAAGAAGCAAGGAGTCAAGGCCCTAGC 1800

Qy  1801 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGAATAATTC 1860
Db  1801 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGAATAATTC 1860

Qy  1861 TTGTTAAGCAAATCTCCTTAAGTAATTATTCAAATAAGATTATACTCATACATCTAT 1920
Db  1861 TTGTTAAGCAAATCTCCTTAAGTAATTATTCAAATAAGATTATACTCATACATCTAT 1920

Qy  1921 ATGTCACTGTTTAAAGAAGATATTTAATTTTAATGTGTGTTACATGGTCTGTAAATATT 1980
Db  1921 ATGTCACTGTTTAAAGAAGATATTTAATTTTAATGTGTGTTACATGGTCTGTAAATATT 1980

Qy  1981 TGTATTTAAAAATGCCATGCCATGCATTAGGCTTTGGAAATTTAGTTGAAATGTAAAAT 2040
Db  1981 TGTATTTAAAAATGCCATGCCATGCATTAGGCTTTGGAAATTTAGTTGAAATGTAAAAT 2040

Qy  2041 GTGAAAAACTTAGATCATTGTAGTAATAAATATTTTTAACTTCATTCATACAGTTAAGT 2100
Db  2041 GTGAAAACTTAGATCATTGTAGTAATAAATATTTTTAACTTCATTCATACAGTTAAGT 2100

Qy  2101 TTATCTGACAATAAAAGCTCTGACTGAATGTTGATTATCCTTCCTATTATGTAATAAGGA 2160

```
Qy        361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT 411

Qy        421 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA 471

Qy        481 TTCAGTGAAGAAAATATTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 TTCAGTGAAGAAAATATTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA 531

Qy        541 CCTCAACAAATTCACCTTAAAGCAAAAGCAATATATGAGAAATTTATACAGACTGATGCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 CCTCAACAAATTCACCTTAAAGCAAAAGCAATATATGAGAAATTTATACAGACTGATGCC 591

Qy        601 CCAAAAGAGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 CCAAAAGAGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAA 651

Qy        661 CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC 711

Qy        721 AGTTATACACGTTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAAGACCTCAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 AGTTATACACGTTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAAGACCTCAG 771

Qy        781 AGACCAACAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 AGACCAACAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA 831

Qy        841 CAATCAGATGTTGCCATTTGGTTATAAAGAAAATTGATTTTGCTCATTTTTATGACAAAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 CAATCAGATGTTGCCATTTGGTTATAAAGAAAATTGATTTTGCTCATTTTTATGACAAAC 891

Qy        901 TTATACATCTGCTTCTAACATATCGCATGTTTATGTTAAGATTTGGTCCCATCCTTTAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        892 TTATACATCTGCTTCTAACATATCGCATGTTTATGTTAAGATTTGGTCCCATCCTTTAAA 951

Qy        961 CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACTTTCTGCTAACAAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACTTTCTGCTAACAAA 1011

Qy       1021 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1012 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTA 1071

Qy       1081 TAATCAGAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAAGTAATAATGTTT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1072 TAATCAGAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAAGTAATAATGTTT 1131

Qy       1141 TATAAGATTGTAGAGTTAAGTAAAAGTTAAGCTTTTGCAAAGTTGTCAAAAGTTCAAACA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1132 TATAAGATTGTAGAGTTAAGTAAAAGTTAAGCTTTTGCAAAGTTGTCAAAAGTTCAAACA 1191

Qy       1201 AAAGTCTAGTTGGGATTTTTTACCAAAGCAGCATAATATGTGTTATATAAACATAATAAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1192 AAAGTCTAGTTGGGATTTTTTACCAAAGCAGCATAATATGTGTTATATAAACATAATAAT 1251

Qy       1261 ACTCAGATATCCAAATGTTCAGATAGCATTTTTCATAATGAATGTTCTCTTTTTTTTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1252 ACTCAGATATCCAAATGTTCAGATAGCATTTTTCATAATGAATGTTCTCTTTTTTTTGGT 1311

Qy       1321 AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1312 AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT 1371

Qy       1381 ACTACTAACCCTGTCCCAAGAATAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 ACTACTAACCCTGTCCCAAGAATAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAA 1431

Qy       1441 CTTTTGTGAAGACACATTCATCTCTACAGAACTTCAGATTAAATATAATCTAGATTAATG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1432 CTTTTGTGAAGACACATTCATCTCTACAGAACTTCAGATTAAATATAATCTAGATTAATG 1491

Qy       1501 ACTGAGAATAAGATCCACATTTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1492 ACTGAGAATAAGATCCACATTTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT 1551

Qy       1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTTGCATATCTCCAGGTAGGGA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1552 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTTGCATATCTCCAGGTAGGGA 1611

Qy       1621 ACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGCCATCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1612 ACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGCCATCA 1671

Qy       1681 ATGGCTTGAATAAAAACCAGAGAAGGTTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1672 ATGGCTTGAATAAAAACCAGAGAAGGTTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG 1731

Qy       1741 AATTGGGGTAGAAATCAGAAATGAGATGAGGGGAAGAAGCAAGGAGTCTAAGGCCCTAGC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1732 AATTGGGGTAGAAATCAGAAATGAGATGAGGGGAAGAAGCAAGGAGTCTAAGGCCCTAGC 1791

Qy       1801 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGATTATATTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1792 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGATTATATTC 1851

Qy       1861 TTGTTAAGCAAATCTCCTTAAGTAATTATTATTCAAATAAGATTATACTCATACATCTAT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1852 TTGTTAAGCAAATCTCCTTAAGTAATTATTATTCAAATAAGATTATACTCATACATCTAT 1911

Qy       1921 ATGTCACTGTTTTAAAGAGATATTTAATTTTTAATGTGTGTTACATGGTCTGTAAATATT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1912 ATGTCACTGTTTTAAAGAGATATTTAATTTTTAATGTGTGTTACATGGTCTGTAAATATT 1971

Qy       1981 TGTATTTAAAAATGCCATGCATTAGGCTTTGGAAATTTAATGTTAGTTGAAATGTAAAAT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1972 TGTATTTAAAAATGCCATGCATTAGGCTTTGGAAATTTAATGTTAGTTGAAATGTAAAAT 2031

Qy       2041 GTGAAAACTTTAGATCATTTGTAGTAATAAATATTTTTAACTTCATTCATACAGTTAAGT 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2032 GTGAAAACTTTAGATCATTTGTAGTAATAAATATTTTTAACTTCATTCATACAGTTAAGT 2091

Qy       2101 TTATCTGACAATAAAAGCTCTGACTGAA 2128
              ||||||||||||||||||||||||||||
Db       2092 TTATCTGACAATAAAAGCTCTGACTGAA 2119

RESULT 5 .
AK057114
LOCUS       AK057114                2136 bp    mRNA    linear   PRI 01-AUG-2002
DEFINITION  Homo sapiens cDNA FLJ32552 fis, clone SPLEN1000089, highly similar
            to Homo sapiens regulator of G-protein signaling 13 mRNA.
ACCESSION   AK057114
VERSION     AK057114.1  GI:16552701
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J.,
            Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K.,
            Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H.,
            Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K.,
            Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2136)
  AUTHORS   Isogai,T., Otsuki,T. and Sugiyama,T.
  TITLE     Direct Submission
```

JOURNAL   Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
           Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
           (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
           Economy, Trade and Industry of Japan; cDNA full insert sequencing:
           Research Association for Biotechnology (RAB); cDNA library
           construction: Helix Research Institute (HRI) (supported by Japan
           Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
           HRI, and Biotechnology Center, National Institute of Technology and
           Evaluation; clone selection for full insert sequencing: RAB and
           HRI.
FEATURES   Location/Qualifiers
     source   1..2136
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /clone="SPLEN1000089"
                   /tissue_type="spleen"
                   /clone_lib="SPLEN1"
                   /note="cloning vector: pME18SFL3"
ORIGIN

Query Match   93.1%;   Score 2064.4;   DB 9;   Length 2136;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 2065;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   61 ATAACTTTTTATTCTACTATGTATATGTATGTGGAATAGTATTAATTAAATGAACTAGGGAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 ATAACTTTTTATTCTACTATGTATATGTATGTGGAATAGTATTAATTAAATGAACTAGGGAAG 130

Qy   121 GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAAGAGATGGAAACACATTGCTTTTC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAAGAGATGGAAACACATTGCTTTTC 190

Qy   181 TTTTCTCAAATAAATATGTGTGAATCAAAAGAAAAGAAAAACTTTTTCAAGTAATACATGGT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   191 TTTTCTCAAATAAATATGTGTGAATCAAAAGAAAAGAAAAACTTTTTCAAGTAATACATGGT 250

Qy   241 TCAGGAAAAGAAGAAGAACAGCAAGAAGAAGAAGCCAAAATCAGAGCTAAGGAAAAAGAAAATAGA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   251 TCAGGAAAAGAAGAAGAACAGCAAGAAGAAGAAGCCAAAATCAGAGCTAAGGAAAAAGAAAATAGA 310

Qy   301 CTAAGTCTTCTTGTGCAGAAACCTGAGTTCATGAAGAACACCCGCTCCAGTAGATCTGGG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   311 CTAAGTCTTCTTGTGCAGAAACCTGAGTTCATGAAGAACACCCGCTCCAGTAGATCTGGG 370

Qy   361 CACTTGGCCAAAGAAACAAGACAGAGTCTCCCTGAGAGGCCAGTGAAATGGGGTGAATCATTT 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   371 CACTTGGCCAAAGAAACAAGACAGAGTCTCCCTGAGAGGCCAGTGAAATGGGGTGAATCATTT 430

Qy   421 GACAAACTGCTTCCATAGAGATGGACTAGAGCCTGGATAGCCTGTGAAGATTTCTTAAAACTGAA 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   431 GACAAACTGCTTCCATAGAGATGGACTAGAGCCTGGATAGCCTGTGAAGATTTCTTAAAACTGAA 490

Qy   481 TTCAGTGAAGAAAATATTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   491 TTCAGTGAAGAAAATATTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA 550

Qy   541 CCTCAACAAATTCACCTTAAAGCAAAGCAAATAATATGAGAAATTTATACAGACTGATGCC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   551 CCTCAACAAATTCACCTTAAAGCAAAGCAAATAATATGAGAAATTTATACAGACTGATGCC 610

Qy   601 CCAAAAGAGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAAACAGCATCACTCAA 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   611 CCAAAAGAGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAAACAGCATCACTCAA 670

Qy   661 CCTACCCTCCACAGTTTGATGTGCACAAAAGCACAGAGTGTATCAGCTTCATGGACAACAGAC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   671 CCTACCCTCCACAGTTTGATGTGCACAAAAGCACAGAGTGTATCAGCTTCATGGACAACAGAC 730

Qy   721 AGTTATACACGTTTCTGAAATCTGACATCTATTAGACTTGATGGAGGAAGACCTCAG 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   731 AGTTATACACGTTTCTGAAATCTGACATCTATTAGACTTGATGGAGGAAGACCTCAG 790

Qy   781 AGACCACAAAATCTTAGGAGACGATCACGCTCATTACCTGCAATGAATTCCAAGATGTA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   791 AGACCACAAAATCTTAGGAGACGATCACGCTCATTACCTGCAATGAATTCCAAGATGTA 850

Qy   841 CAATCAGATGTGCCATTTGGTTATAAAGAAAATTGATTTTGCTCATTTTATGACAAAC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   851 CAATCAGATGTGCCATTTGGTTATAAAGAAAATTGATTTTGCTCATTTTATGACAAAC 910

Qy   901 TTATACATCTGCTTCTAACATATCGCATGTTTATGTTAAGATTGGTCCCATCCTTTAAA 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   911 TTATACATCTGCTTCTAACATATCGCATGTTTATGTTAAGATTGGTCCCATCCTTTAAA 970

Qy   961 CTGAAATATGTCATGTGAAATATTTTAAAAATGTAAAAACAAACTTTCTGCTAACAAA 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   971 CTGAAATATGTCATGTGAAATATTTTAAAAATGTAAAAACAAACTTTCTGCTAACAAA 1030

Qy 1021 ATACATACAGTATCTGCCAGTATATCTGTAAAACCTTCTATTTGATGTCATTCCATTTA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1031 ATACATACAGTATCTGCCAGTATATCTGTAAAACCTTCTATTTGATGTCATTCCATTTA 1090

Qy 1081 TAATCAGAAAAAAAAAACTTATTCTTAATCAAAAGGCAGTACAAAAAAGTAATAATGTTT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1091 TAATCAGAAAAAAAAAACTTATTCTTAATCAAAAGGCAGTACAAAAAAGTAATAATGTTT 1150

Qy 1141 TATAAGATTGTAGAGTTAAGTAAAAGTTAAGTAAAAGTTTGCAAAGTGTCAAAAGTCAAACA 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1151 TATAAGATTGTAGAGTTAAGTAAAAGTTAAGTAAAAGTTTGCAAAGTGTCAAAAGTCAAACA 1210

Qy 1201 AAAGTCTAGTGGGATTTTTTACCAAAGCAGCAGATAATATGTGTATATAAACATAATAAT 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1211 AAAGTCTAGTGGGATTTTTTACCAAAGCAGCAGATAATATGTGTATATAAACATAATAAT 1270

Qy 1261 ACTCAGATATCCAAATGTTCAGATAGCATTTTTCATAATGATGTGTTCTCTTTTTTTTTGGT 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1271 ACTCAGATATCCAAATGTTCAGATAGCATTTTTCATAATGATGTGTTCTCTTTTTTTTTGGT 1330

Qy 1321 AATAGTGTAGAAGTGATCGGTTCTTACAATGGGAGAATGAAGAACAACATTTATTGGTT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1331 AATAGTGTAGAAGTGATCGGTTCTTACAATGGGAGAATGAAGAACAACATTTATTGGTT 1390

Qy 1381 ACTACTAACCCTGTCCCAAGACACATTCATTCTCTACAGAATGTATATCTACAGAACTAATG 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1391 ACTACTAACCCTGTCCCAAGACACATTCATTCTCTACAGAATGTATATCTACAGAACTAATG 1450

Qy 1441 CTTTTGTGAAGACACATTCATTCTCACTTGAACTCCTCAAGTGAACATGGACGTACCCAGTTAT 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1451 CTTTTGTGAAGACACATTCATTCTCACTTGAACTCCTCAAGTGAACATGGACGTACCCAGTTAT 1510

Qy 1501 ACTGAGAATAAGATCCACATTCACCAGAAGGTTTTCCCAGAGAAGCAAGGAGTCAAGGTAGGGA 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1511 ACTGAGAATAAGATCCACATTCACCAGAAGGTTTTCCCAGAGAAGCAAGGAGTCAAGGTAGGGA 1570

Qy 1561 ACAAAGTACTTCTGTTGGTCACAGAAACAATGAGATGAGGGGGAAGCAAGAGTGTTATCTCCAGGTAGGGA 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1571 ACAAAGTACTTCTGTTGGTCACAGAAACAATGAGATGAGGGGGAAGCAAGAGTGTTATCTCCAGGTAGGGA 1630

Qy 1621 ACTAAGTAGACTACCTTATCACCGGCTAAGAAACTTGCTACTAAACTATTAGGCCATCA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1631 ACTAAGTAGACTACCTTATCACCGGCTAAGAAACTTGCTACTAAACTATTAGGCCATCA 1690

Qy 1681 ATGGCTTGAATAAAACCAGAGAAGGTTTTCCCAGAGCGTCCATGTTTGGCCCTTTAG 1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1691 ATGGCTTGAATAAAACCAGAGAAGGTTTTCCCAGAGCGTCCATGTTTGGCCCTTTAG 1750

Qy 1741 AATTGGGGTAGAAATCAGAAATGAGAATGAGATGGGGAAGCAAGGAGTCTAAGGCCCTAGC 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1751 AATTGGGGTAGAAATCAGAAATGAGAATGAGATGGGGAAGCAAGGAGTCTAAGGCCCTAGC 1810

Qy 1801 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAAGTGTTATCTCATTGATTATATTC 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1811 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAAGTGTTATCTCATTGATTATATTC 1870

```
Qy   1861  TTGTTAAGCAAATCTCCTTAAGTAATTATTATTATTCAAATAAGATTATACTCATACATCTAT  1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1871  TTGTTAAGCAAATCTCCTTAAGTAATTATTATTATTCAAATAAGATTATACTCATACATCTAT  1930

Qy   1921  ATGTCACTGTTTAAAGAGATATTTAATTTTTAATGTGTTACATGTCTGTAAATATT  1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1931  ATGTCACTGTTTAAAGAGAGATATTTAATTTTTAATGTGTTACATGGTCTGTAAATATT  1990

Qy   1981  TGTATTTAAAAATGCCATGCATTAGGCTTTGGAAATTAATGTTAGTTGAAATGTAAAAT  2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1991  TGTATTTAAAAATGCCATGCATTAGGCTTTGGAAATTAATGTTAGTTGAAATGTAAAAT  2050

Qy   2041  GTGAAAACTTAGATCATTTGTAGTAATAAAATATTTTAACTTCATTCATACAGTTAAGT  2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2051  GTGAAAACTTTAGATCATTTGTAGTAATAAAATATTTTTAACTTCATTCATACAGTTAAGT  2110

Qy   2101  TTATCTGACAATAAAAGCTCTGACTG  2126
           ||||||||||||||||||||||||||
Db   2111  TTATCTGACAATAAAAGCTCTGACTG  2136
```

RESULT 6
BC020632
LOCUS   BC020632   2164 bp   mRNA   linear   PRI 04-OCT-2003
DEFINITION   Homo sapiens regulator of G-protein signalling 18,   mRNA (cDNA clone
             MGC:22266 IMAGE:4714909), complete cds.
ACCESSION   BC020632
VERSION   BC020632.1   GI:18088322
KEYWORDS   MGC.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2164)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
             Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
             Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
             Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
             Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
             Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
             Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
             Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
             Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
             McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
             Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
             Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
             Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
             Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
             Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
             Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
             Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
             Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-length
             human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2 (bases 1 to 2164)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (03-JAN-2002) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National.Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
  REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
  COMMENT   Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov
             Tissue procurement: CLONTECH
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Sequencing Group at the Stanford Human Genome
             Center, Stanford University School of Medicine, Stanford, CA 94305
             Web site: http://www-shgc.stanford.edu
```

```
             Contact: (Dickson, Mark) mcd@paxil.stanford.edu
             Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
             R. M.

             Clone distribution: MGC clone distribution information can be found
             through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
             Series: IRAL Plate: 37 Row: d Column: 18
             This clone was selected for full length sequencing because it
             passed the following selection.criteria: matched mRNA gi: 18640749.
FEATURES             Location/Qualifiers
     source          1..2164
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:22266 IMAGE:4714909"
                     /tissue_type="Liver"
                     /clone_lib="NIH_MGC_76"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
     gene            1..2164
                     /gene="RGS18"
                     /note="synonym: RGS13"
                     /db_xref="LocusID:64407"
                     /db_xref="MIM:607192"
     CDS             169..876
                     /codon_start=1
                     /product="regulator of G-protein signalling 18"
                     /protein_id="AAH20632.1"
                     /db_xref="GI:18088323"
                     /db_xref="LocusID:64407"
                     /translation="METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEK
                     RNRLSLLVQKPEFHEDTRSSRSGHLAKSTRVSPEEAVKWGESFDKLLSHRDGLEAFTR
                     FLKTEFSEENIEFWIACEDFKKSKGPQOIHLKAIYEKFIQTDAPKEVNLDFHTKEV
                     ITNSITQPTLHSFDAAQSRVYQLMEQDSTRFLKSDIYLDLMEGRPQRPTNLRRRSRS
                     FTCNEFQDVQSDVAIWL"
     misc_feature    424..771
                     /note="RGS; Region: Regulator of G protein signalling
                     domain"
                     /db_xref="CDD:smart00315"
ORIGIN
     Query Match             93.1%;  Score 2063.2;  DB 9;  Length 2164;
     Best Local Similarity   99.9%;  Pred. No. 0;
     Matches 2065;  Conservative   0;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy    61  ATAACTTTTATTCTACTATGTATATGTGTGGAATAGTATTAATAATAAATGAACTAGGGAAG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    70  ATAACTTTTATTCTACTATGTATATGTGTGGAATAGTATTAATAATAAATGAACTAGGGAAG  129

Qy   121  GATGTAATAAATTAGACATCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   130  GATGTAATAAATTAGACATCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC  189

Qy   181  TTTTCCAAATAAATATGTGAATCAAAGAGGAAATCAGAGCTAAGGAAAAATCAGAAATGGT  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   190  TTTTCCAAATAAATATGTGAATCAAAGAGGAAATCAGAGCTAAGGAAAAATCAGAAATGGT  249

Qy   241  TCAGGAAAGAAGAAGAAACAAGCAAAGCAAAGAGAGCCAAAATCAGAGCTAAGGAAAAATAGA  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   250  TCAGGAAAGAAGAAGAAACAAGCAAAGCAAAGAGAGCCAAAATCAGAGCTAAGGAAAAATAGA  309

Qy   301  CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAGAAGATCTCGGG  360
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   310  CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAGAAGATCTCGGG  369

Qy   361  CACTTGGCCAAGAAACAAGAGTCTCCCCTGAAGAGCAGTGAATGGGTGAATCATTT  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   370  CACTTGGCCAAGAAACAAGAGTCTCCCCTGAAGAGCAGTGAAATGGGTGAATCATTT  429

Qy   421  GACAAACTGCTTCCCATAGAGATGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   430  GACAAACTGCTTCCCATAGAGATGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA  489
```

```
Qy   481 TTCAGTGAAGAAATATTTGGATAGCCTGTGAAGATTCAAGAAAGCAAGGGA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   490 TTCAGTGAAGAAATATTTGGATAGCCTGTGAAGATTCAAGAAAGCAAGGGA 549

Qy   541 CCTCAACAAATTCACCTTAAAGCAAAAGCAATATGAGAAATTTATACAGACTGATGCC 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   550 CCTCAACAAATTCACCTTAAAGCAAAAGCAATATGAGAAATTTATACAGACTGATGCC 609

Qy   601 CCAAAAGAGGTAACCTGATTTCACACAAAGAGAGTCATTACAACAGCATCACTCAA 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   610 CCAAAAGAGGTAACCTGATTTCACACAAAGAGAGTCATTACAACAGCATCACTCAA 669

Qy   661 CCTACCCTCCACAGTTTGATGTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   670 CCTACCCTCCACAGTTTGATGTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC 729

Qy   721 AGTTATACACGTTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAAGACCTCAG 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   730 AGTTATACACGTTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAAGACCTCAG 789

Qy   781 AGACCAACAAATCTTAGGAGACGAGCATCACGCTCATTTAGACTTGCTCATTTTATGACAAAC 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   790 AGACCAACAAATCTTAGGAGACGAGCATCACGCTCATTTAGACTTGCTCATTTTATGACAAAC 849

Qy   841 CAATCAGATGTGTTCCATTGGTATAAAGAAAATTGATTTTGCTCATTTTATGTCCATCCTTTAAA 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   850 CAATCAGACGTTGCCATTGGTATAAAGAAAATTGATTTTGCTCATTTTATGTCCATCCTTTAAA 909

Qy   901 TTATACACTCTGCTTCTCAACATATCGCATGTTTATGTGTTAAGATTGGTCCATCCTTTAAA 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   910 TTATACACTCTGCTTCTCAACATATCGCATGTTTATGTGTTAAGATTGGTCCATCCTTTAAA 969

Qy   961 CTGAAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACCTTCTGCTAACAA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   970 CTGAAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACCTTCTGCTAACAA 1029

Qy  1021 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTGTCATTCATTTA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1030 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCATTTA 1089

Qy  1081 TAATCAGAAAAAAACTTATTCTTAATCAAAAGGCAGTACAAAAAAAAGTAATAAATGTTT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1090 TAATCAGAAAAAAACTTATTCTTAATCAAAAGGCAGTACAAAAAAAAGTAATAAATGTTT 1149

Qy  1141 TATAAGATTGTAGAGTTAAGTAAAAGCTTTGCAAAGTTCAAAAGTTCAAACA 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1150 TATAAGATTGTAGAGTTAAGTAAAAGCTTTGCAAAGTTGTCAAAAGTTCAAACA 1209

Qy  1201 AAAGTCTAGTTGGGGATTTTTACCAAAGCAGCATAAATGTGTTATAAAACATAATAAT 1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1210 AAAGTCTAGTTGGGGATTTTTACCAAAGCAGCATAAATGTGTTATAAAACATAATAAT 1269

Qy  1261 ACTCCAGATATCCAAAGTGTCAGATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTTTTGGT 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1270 ACTCCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTTTTGGT 1329

Qy  1321 AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1330 AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT 1389

Qy  1381 ACTACTAACCCTGTCCCAAGAATAGTAATATCACCTCTAGTTATGTATAAGCCAGCAACAGGAA 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1390 ACTACTAACCCTGTCCCAAGAATAGTAATATCACCTCTAGTTATGTATAAGCCAGCAACAGGAA 1449

Qy  1441 CTTTTGTGAAGACACATTCATCTCTACGAACTTCAGATTAAATATAATATCTAGATTAATG 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1450 CTTTTGTGAAGACACATTCATCTCTACGGAACTTCAGATTAAATATAATATCTAGATTAATG 1509

Qy  1501 ACTGAGAATAAGATCCACATTTGAACTCATTCCTAAGTGACGTACCCAGTTAT 1560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1510 ACTGAGAATAAGATCCACATTCCAACTCATTCCTAAGTGACGTACCCAGTTAT 1569

Qy  1561 ACAAAGTACTTCTGTTCGTCACAGAAACATGACCAGATTTTGCATATCTCCAGGTAGGGA 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1570 ACAAGTACTTCTGTTCGTCACAGAAACATGACCAGATTTTGCATATCTCCAGGTAGGGA 1629

Qy  1621 ACTAAGTAGACTACCTATCACCGGCTAAGAAACCTGCTACTAAACTATTAGGCCATCA 1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1630 ACTAAGTAGACTACCTATCACCGGCTAAGAAACCTGCTACTAAACTATTAGGCCATCA 1689

Qy  1681 ATGGCTTGAATAAAAACCAGAGAAGTTTTCCCAGGACGTCTCATGTTTGGCCCCTTTAG 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1690 ATGGCTTGAATAAAAACCAGAGAAGGTTTTCCCAGGACGTCTCATGTTTGGCCCCTTTAG 1749

Qy  1741 AATTGGGTAGAAATCAGAAATGAGATGAGGGGAAGCAAGGAGTCTAAGGCCCTAGC 1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1750 AATTGGGTAGAAATCAGAAATGAGATGAGGGGAAGCAAGGAGTCTAAGGCCCTAGC 1809

Qy  1801 GATTTGGGCATCTGCCACCATTGGTTCATATTCAGAAAGTGTATCTCATTGATTATATTC 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1810 GATTTGGGCATCTGCCACCATTGGTTCATATTCAGAAAGTGTATCTCATTGATTATATTC 1869

Qy  1861 TTGTTAAGCAAATCTCCTTAAGTAATTATTATTCAAATAAGATTATACTCATACATCTAT 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1870 TTGTTAAGCAAATCTCCTTAAGTAATTATTATTCAAATAAGATTATACTCATACATCTAT 1929

Qy  1921 ATGTCACTGTTTTAAAGAGATATTTAATTTTAATTTTTAATGTGTGTTACATGGTCTGTAAATATT 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1930 ATGTCACTGTTTTAAAGAGATATTTAATTTTAATTTTTAATGTGTGTTACATGGTCTGTAAATATT 1989

Qy  1981 TGTATTTAAAAATGCCATGCATTAGGCTTGGAAATTTAATGTTAGTTGAAATGTAAAAT 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1990 TGTATTTAAAAATGCCATGCATTAGGCTTGGAAATTTAATGTTAGTTGAAATGTAAAAT 2049

Qy  2041 GTGAAAACTTAGATCATTGTAGTAATAAACATATTTTAACTTCATTCATACAGTTAAGT 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2050 GTGAAAACTTAGATCATTGTAGTAATAAACATATTTTAACTTCATTCATACAGTTAAGT 2109

Qy  2101 TTATCTGACAATAAAAGCTCTGACTGAA 2128
         |||||||||||||||||||||||||||
Db  2110 TTATCTGACAATAAAAGCTCTGACTAA 2137
```

RESULT 7
AX299976

```
LOCUS       AX299976                2144 bp    DNA     linear   PAT 26-NOV-2001
DEFINITION  Sequence 19 from Patent WO0183514.
ACCESSION   AX299976
VERSION     AX299976.1  GI:17129463
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Murray,D.L. and Gagnon,A.W.
  TITLE     Nucleic acids encoding a novel regulator of G protein signaling,
            rgs18, and uses thereof
  JOURNAL   Patent: WO 0183514-A19 08-NOV-2001;
            Aventis Pharmaceuticals Products Inc. (US)
FEATURES    Location/Qualifiers
     source 1..2144
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
```

```
Query Match            92.1%;  Score 2041.6;  DB 6;  Length 2144;
Best Local Similarity  99.5%;  Pred. No. 0;
Matches 2058;  Conservative 0;  Mismatches 9;  Indels 1;  Gaps 1;

Qy    61 ATAACTTTTATTCTACTATGTATATGTGTAATAATGATATAATAATGACTAGGGAAG 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    64 ATAACTTTTATTCTACTATGTATATGTGTAATAATGATATAATAATGACTAGGGAAG 123

Qy   121 GATGTAATAATTAGACATCTTCATTTAGAGAGAAGAGAATGGAAACAACATTGCTTTTC 180
```

Mon Aug 23 07:18:08 2004

us-09-894-749-1.rge

```
Db          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1204  AAAGTCTAGTTGGGATTTTTTACCAAAGCAGCATAATATGTGTTATATAAACATAATAATAAT  1263
Qy    1261  ACTCAGATAATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTTTTGGT  1320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1264  ACTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTTTTGGT  1323
Qy    1321  AATAGTGTAGAGTGATCTCGGTTCTTACAATGGGAGATGAAGAACATTATTATTGGGTT  1380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1324  AATAGTGTAGAGTGATCTCGGTTCTTACAATGGGAGATGAAGAACATTATTATTGGGTT  1383
Qy    1381  ACTACTAACCCTGTCCCAAGAATAGTAATATCCTAGTTATAAGCCAGCAACAGGAA  1440
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1384  ACTACTAACCCTGTCCCAAGAATAGTAATATCCTAGTTATAAGCCAGCAACAGGAA  1443
Qy    1441  CTTTTGTGAAGACACATTCATCTTCACAGAACTTCAGATTAAATATAATCTAGATTAATG  1500
            |||||||| |||||||||||||||||||||||||||||| |||||||| |||||||||
Db    1444  CTTTTGTGAAGACACACATTCATCTTCACAGAACTTCAGATTAAATATAATCTAGATTAATG  1503
Qy    1501  ACTGAGAATAAGATCCACATTGAACTCATCCTAAGTGAACATGGACGTACCCAGTTAT  1560
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1504  ACTGAGAATAAGATCCACATTGAACTCATCCTAAGTGAACATGGACGTACCCAGTTAT  1563
Qy    1561  ACAAAGTACTTCTGTTGGTCACAGAAACTTGCATATCTCCAGGTAGGGA  1620
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db    1564  ACAAAGTACTTCTGTTGGTCACAGAAACACTGAGAAACATGCATATCTCCAGGTAGGGA  1623
Qy    1621  ACTAAGTAGACTACCTTATCACCGGCTAAGAAGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG  1680
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1624  ACTAAGTAGACTACCTTATCACCGGCTAAGAAGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG  1683
Qy    1681  ATGGCTTGAATAAAAACCAGAGAAGTGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG  1740
            |||||||||||||||||||| ||||| ||||||||||||||||||||||||||||
Db    1684  ATGGGCTGGAATAAAAACCGAGAA-GTTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG  1742
Qy    1741  AATTGGGTAGAAATCAGAAATGAGATGAGGGAGAAGCAAGGAGTCTAAGGCCCTAGC  1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1743  AATTGGGTAGAAATCAGAAATGAGATGAGGGAGAAGCAAGGAGTCTAAGGCCCTAGC  1802
Qy    1801  GATTTGGGCATCTGCACACTGGTTCATATTCAGAAAGTGTTATCTCATTGATTATATTC  1860
            |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    1803  GATTTGGGCATCTGCACACTGGTTCATATTCAGAAAGTGTTATCTCATTGATTATATTC  1862
Qy    1861  TTGTTAAGCAAATCTCCTTAAGTAATTATTATTATTATTTCAAATAAGATTATACTCATCATCTAT  1920
            ||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    1863  TTGTTAAGCAAATCTCCTTAAGTAATTATTATTATTTCAAATAAGATTATACTCATCATCTAT  1922
Qy    1921  ATGTCACTGTTTTAAAGAGATATTTAATTTTTAATTGTGTTACATGGTCTGTAAATATT  1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1923  ATGTCACTGTTTTAAAGAGATATTTAATTTTTAATTGTGTTACATGGTCTGTAAATATT  1982
Qy    1981  TGTATTTAAAAATGCCATGCATTAGGCTTTGGAAATTAATGTTAGTTGAAATGTAAAAT  2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1983  TGTATTTAAAAATGCCATGCATTAGGCTTTGGAAATTAATGTTAGTTGAAATGTAAAAT  2042
Qy    2041  GTGAAAACTTTAGATCATTGTAGTAATAAATATTTTAACTTCATTCATTACACAGTTAAGT  2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2043  GTGAAAACTTTAGATCATTGTAGTAATAAATATTTTAACTTCATTCATACAATTAAGT  2102
Qy    2101  TTATCTGACAATAAAAGCTCTGACTGAA  2128
            ||||||||||||||||||||||||||||
Db    2103  TTATCTGACAATAAAAGCTCTGACTGAA  2130
```

```
Db           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       124   GATGTAATAAATTAGACATCTCTTCATTTAGAGGAGAGATGATGGAAACAACAACAACTTGCTTTTC  183
Qy     181   TTTTCTCAAATAAATATGTGAATCAAAAGAAAAAACTTTTTTCAAGTTAATACATGGT  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     184   TTTTCTCAAATAAATATGTGAATCAAAAGAAAAAACTTTTTTCAAGTTAATACATGGT  243
Qy     241   TCAGGAAAGAAGAAACAAGAGCCAAAATCAGAGCTAAGGAAAAAGAGAAATAGA  300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     244   TCAGGAAAGAAGAAACAAGAGCCAAAATCAGAGCTAAGGAAAAAGAGAAATAGA  303
Qy     301   CTAAGTCTTCTTGTGCGAAAACCTGAGTTTCATGAAGACACACCCGCTCCAGTAGATCTGGG  360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     304   CTAAGTCTTCTTGTGCGAAAACCTGAGTTTCATGAAGACACACCCGCTCCAGTAGATCTGGG  363
Qy     361   CACTTGGCCAAAGAACAAGAGTCTCCCCTGAAGAGCAGTCAGTGAATGGGTGTGAATCATTT  420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     364   CACTTGGCCAAAGAACAAGAGTCTCCCCTGAAGAGCAGTCAGTGAATGGGTGTGAATCATTT  423
Qy     421   GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTACCAGATTTCTTAAAACTGAA  480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     424   GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTACCAGATTTCTTAAAACTGAA  483
Qy     481   TTCAGTGAAGAAAATATTTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAAAGCAAGGGA  540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     484   TTCAGTGAAGAAAATATTTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAAAGCAAGGGA  543
Qy     541   CCTCAACAAAATTCACCCTAAAGCAAAGCAATATGAAATTTATACAGAGCTGATGCC  600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     544   CCTCAACAAAATTCACCCTAAAGCAAAGCAATATGAAATTTATACAGAGCTGATGCC  603
Qy     601   CCAAAAGAGGTTAACCTTGATTTTCACCACAAAGAAGTCATTACAAAACAGCACTCACTCAA  660
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     604   CCAAAAGAGGTTAACCTCGATTTTCACCACAAAGAAGTCATTACAAAACAGCACTCACTCAA  663
Qy     661   CCTACCCTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCAGGAACAAGAC  720
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     664   CCTACCCTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCAGGAACAAGAC  723
Qy     721   AGTTATACACGTTTTCTGAAATCTATTCTGACATCTATTAGACTTGATGGAAGGAAGACCTCAG  780
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     724   AGTTATACACGTTTTCTGAAATCTATTCTGACATCTATTAGACTTGATGGAAGGAAGACCTCAG  783
Qy     781   AGACCACAACAAATCTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA  840
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     784   AGACCACAACAAATCTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA  843
Qy     841   CAATCAGATGTTGCATGTGAAAATATTTTAAAATGTAAAAACAAAACTTTTGTGACAAAC  900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     844   CAATCAGATGTTGCATGTGAAAATATTTTAAAATGTAAAAACAAAACTTTTGTGACAAAC  903
Qy     901   TTTAACATCTGCTTCTAACATGGCATGTTATGTTAAGATTTGGTCCCATCCTTTAAA  960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     904   TTTAACATCTGCTTCTAACATGGCATGTTATGTTAAGATTTGGTCCCATCCTTTAAA  963
Qy     961   CTGAAATATGTCATGTGAAATATTTTAAAATGTAAAAACAAAACTTTTCTGCTAACAAA  1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     964   CTGAAATATGTCATGTGAAATATTTTAAAATGTAAAAACAAAACTTTTCTGCTAACAAA  1023
Qy    1021   ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTA  1080
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1024   ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTA  1083
Qy    1081   TAATCAGAGAAAAAAAAACTTATTCTTAATCAAAAGCAGTACAAAAAAGTAATAAATGTTT  1140
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1084   TAATCAGAGAAAAAAAAACTTATTCTTAATCAAAAGCAGTACAAAAAAGTAATAAATGTTT  1143
Qy    1141   TATAGATTGTGAGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTTGTCAAAACA  1200
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1144   TATAGATTGTGAGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTTGTCAAAACA  1203
Qy    1201   AAAGTCTAGTTGGGATTTTTTACCAAAGCAGCATAATATGTGTTATATAAACATAATAAT  1260
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2144)
AUTHORS     Gagnon,A.W., Murray,D.L. and Leadley,R.J. Jr.
TITLE       Cloning and characterization of a novel regulator of G protein
            signaling in human platelets
JOURNAL     Cell. Signal. 14 (7), 595-606 (2002)
MEDLINE     21952499
PUBMED      11955952
REFERENCE   2 (bases 1 to 2144)
AUTHORS     Gagnon,A.W., Murray,D.L. and Leadley,R.J. Jr.
TITLE       Direct Submission
JOURNAL     Submitted (16-MAY-2000) Pfizer Central Research, Eastern Point Rd,
            Groton, CT 06340, USA
FEATURES             Location/Qualifiers
     source          1..2144
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="platelet"
     CDS             163..870
                     /codon_start=1
                     /product="regulator of G-protein signaling 18"
                     /protein_id="AAK58589.1"
                     /db_xref="GI:14279411"
                     /translation="METTLLFFSQINMCESKEKTFPKLIHGSGKEETSKEAKIRAKEK
                     RNRLSLLVQKPEHEDTRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTR
                     FLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEV
                     ITNSITQPTLHSFDAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRS
                     FTCNEFQDVQSDVAIWL"
ORIGIN
```

```
Query Match       92.1%;   Score 2041.6;  DB 9;  Length 2144;
Best Local Similarity   99.5%;   Pred. No. 0;
Matches 2058;   Conservative_  0;  Mismatches  9;  Indels  1;  Gaps  1;

Qy     61  ATAACTTTTTATTCTACTATGTATATGTATATGGAATAGTATTAATAAATGAACTAGGGAAG   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64  ATAACTTTTTATTCTACTATGTATATGTATATGGAATAGTATTAATAAATGAACTAGGGAAG   123

Qy    121  GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAGAATGAAACACATTGCTTTTC   180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124  GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAGAATGAAACACATTGCTTTTC   183

Qy    181  TTTTCTCAAATAATAATGTGAATCAAAAGAGAAAAAACTTTTTCAGTTAATACATGGT   240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184  TTTTCTCAAATAATAATGTGAATCAAAAGAGAAAAAACTTTTTCAGTTAATACATGGT   243

Qy    241  TCAGGAAAAGAAGAAGCAAAGAAGCCAAAATCAGAGCTCAAGAAGAAAAAAGAAATAGA   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244  TCAGGAAAAGAAGAAGCAAAGAAGCCAAAATCAGAGCTCAAGAAGAAAAAAGAAATAGA   303

Qy    301  CTAAGTCTTCTGTGCAGAAACCTGAGTTCATGAAGACACCCGCTCCAGTAGATCTGGG   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304  CTAAGTCTTCTGTGCAGAAACCTGAGTTCATGAAGACACCCGCTCCAGTAGATCTGGG   363

Qy    361  CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTT   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364  CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTT   423

Qy    421  GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424  GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA   483

Qy    481  TTCAGTGAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA   540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484  TTCAGTGAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA   543

Qy    541  CCTCAACAAATTCACCTTAAAGCAAAAGCAATTATATGAGAAAATTATATACAGACTGATGCC   600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544  CCTCAACAAATTCACCTTAAAGCAAAAGCAATTATATGAGAAAATTATATACAGACTGATGCC   603

Qy    601  CCAAAAGAGGTTAACCTTTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAA   660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604  CCAAAAGAGGTTAACCTTCGATTTTCACACAAAAGAAGTCATTACAACAGCATCACTCAA   663

Qy    661  CCTACCCTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664  CCTACCCTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC   723

Qy    721  AGTTATACACGTTTCTGAAATCTGAAATCTGACATCATTAGACTTGATGGAAGGAAGACCTCAG   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    724  AGTTATACACGTTTCTGAAATCTGAAATCTGACATCATTAGACTTGATGGAAGGAAGACCTCAG   783

Qy    781  AGACCAACAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA   840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    784  AGACCAACAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA   843

Qy    841  CAATCAGATGTGCCATTGGTTATAAAGAAAATTGATTTTGCTCATTTTATGACAAAC   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844  CAATCAGATGTGCCATTGGTTATAAAGAAAATTGATTTTGCTCATTTTATGACAAAC   903

Qy    901  TTATACATCTGCTTCAACATATCGATGTTTATGTTAAGAATTGGTCCCATCCTTTAAA   960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904  TTATACATCTGCTTCAACATATCGATGTTTATGTTAAGAATTGGTCCCATCCTTTAAA   963

Qy    961  CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAACTTTCTGCTAACAAA   1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    964  CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAACTTTCTGCTAACAAA   1023

Qy   1021  ATACATACAGATCTGCCAGTATATTCGTAAAACCTTCTATTTGATGTCATTCCATTTA   1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1024  ATACATACAGATCTGCCAGTATATTCGTAAAACCTTCTATTTGATGTCATTCCATTTA   1083

Qy   1081  TAATCAGAAAAAAACTTATTCTTATCAAAAGGCGTACAAAAAGTAATAAATGTTT   1140
           ||||||||||||||||||||||| |||||||||||| |||||||||||||||||||
Db   1084  TAATCAGAAAAAAACTTATTCTTAATCAAAAGGCAGTACAAAAAGTAATAAATGTTT   1143

Qy   1141  TATAAGATTGTAGAGTTAAGTAAAGTTAAGCTTTTGCAAAGTTGTCAAAAGTTCAAACA   1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1144  TATAAGATTGTAGAGTTAAGTAAAGTTAAGCTTTTGCAAAGTTGTCAAAAGTTCAAACA   1203

Qy   1201  AAAGTCTAGTGGGATTTTTACCAAGCAGCATAATATGTGTATATAAACATAATAAT   1260
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1204  AAAGTCTAGTGGGATTTTTACCAAGCAGCATAATATGTGTATAAATAAACATAATAAT   1263

Qy   1261  ACTCAGATATCCAAATGTTCAGTTAATTTTTCATGTTAATACATGGT   1320
           ||||||||||||||||||||||||||||||||| |||||||||||||
Db   1264  ACTCAGATATCCAAATGTTCAGTTAATTTTTCATAATGAATGTTTCATAATACATGGT   1323

Qy   1321  AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGACGATGAAGAACATTTATTATTGGGTT   1380
           ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1324  AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGGATGAAGAACATTTATTATTGGGTT   1383

Qy   1381  ACTACTAACCCTGTCCCAAGAATAGTAATATCACCTCAGTATAGAGCCAGCAACAGGAA   1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1384  ACTACTAACCCTGTCCCAAGAATAGTAATATCACCTCAGTATAGAGCCAGCAACAGGAA   1443

Qy   1441  CTTTTGTGAAGACACAATTCATCTCACAGAACTTCAGATTAAATAATCTAGATTAATG   1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1444  CTTTTGTGAAGACACAATTCATCTCACAGAACTTCAGATTAAATAATCTAGATTAATG   1503

Qy   1501  ACTGAGAATAAGATCCATTGAACTCATTCCTAAGTGACATGAACGTACCCAGTTAT   1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1504  ACTGAGAATAAGATCCATTGAACTCATTCCTAAGTGACATGAACGTACCCAGTTAT   1563

Qy   1561  ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTTGCATATCTCCAGGTAGGGA   1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1564  ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTTGCATATCTCCAGGTAGGGA   1623

Qy   1621  ACTAAGTAGACTACCTTATCACCGGCTAAGAAAAACTTGCTACTAAACTATTAGGCCATCA   1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1624  ACTAAGTAGACTACCTTATCACCGGCTAAGAAAAACTTGCTACTAAACTATTAGGCCATCA   1683

Qy   1681  ATGGCTTGAATAAAAACCAGAGAAGGTTTTCCCGGACGTCTCATGTTTGGCCCTTAG   1740
           |||| ||   |||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1684 ATGGGCTCGAATAAAAACCGAGAA-GTTTTTCCCAGGACGTCTCATGTGTTGGCCCTTTAG 1742
Qy   1741 AATTGGGTAGTAGAAATCAGAAATGAGATGAGGGGAAGAGCAAGGAGTCTAAGGCCCTAGC 1800
          |||||||                                            ||||||||
Db   1743 AATTGGGTAGTAGAAATCAGAAATGAGATGAGGGGAAGAGCAAGGAGTCTAAGGCCCTAGC 1802

Qy   1801 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGATTATATTC 1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1803 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGATTATATTC 1862

Qy   1861 TTGTTAAGCAAATCTCCTTAAGTAATAATTATTATTCAAATAAGATTATACTCATCACATCTAT 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1863 TTGTTAAGCAAATCTCCTTAAGTAATAATTATTATTCAAATAAGATTATACTCATCACATCTAT 1922

Qy   1921 ATGTCACTGTGTTTAAAAGAGATATTTAATTTTAATGTGTGTTACATGGTCTGTAAATATT 1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1923 ATGTCACTGTGTTTAAAAGAGATATTTAATTTTAATGTGTGTTACATGGTCTGTAAATATT 1982

Qy   1981 TGTATTTAAAAATGCCATGCCATTAGGCTTTTGGAAATTTAGTTGGAAATGTAAAAT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1983 TGTATTTAAAAATGCCATGCCATTAGGCTTTTGGAAATTTAGTTGGAAATGTAAAAT 2042

Qy   2041 GTGAAAACTTTAGATCATCATTTGTAGTAATAAATAAATATTTTAACCTTCATTCATTACAGTTAAGT 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2043 GTGAAAACTTTAGATCATCATTTGTAGTAATAAATAAATATTTTAACCTTCATTCATTACAGTTAAGT 2102

Qy   2101 TTATCTGACAATAAAAGCTCTGACTGAA 2128
          ||||||||||||||||
Db   2103 TTATCTGACAATAAAAGCTCTGACTGAA 2130
```

RESULT 9
AX299975

```
LOCUS       AX299975                   1840 bp    DNA        linear   PAT 26-NOV-2001
DEFINITION  Sequence 18 from Patent WO0183514.
ACCESSION   AX299975
VERSION     AX299975.1   GI:17129462
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Murray,D.L. and Gagnon,A.W.
  TITLE     Nucleic acids encoding a novel regulator of G protein signaling,
            rgs18, and uses thereof
  JOURNAL   Patent: WO 0183514-A 18 08-NOV-2001;
            Aventis Pharmaceuticals Products Inc. (US)
FEATURES             Location/Qualifiers
     source          1..1840
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
Query Match             79.1%;  Score 1753.2;  DB 6;  Length 1840;
Best Local Similarity   99.5%;  Pred..No. 1.1e-275;
Matches 1769;  Conservative   0;  Mismatches   8;  Indels   1;  Gaps   1;

Qy     61 ATAACTTTTTATTCTACTACTATGTATATGTATGGAATAATAGTAATAAATGAACTAGGGAAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 ATAACTTTTTATTCTACTACTATGTATATGTATGGAATAATAGTAATAAATGAACTAGGGAAG 123

Qy    121 GATGTAATAAATTAGAACATCTTCATTTAGAGAAGAAGATGGAAACAACATTGCTTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 GATGTAATAAATTAGAACATCTTCATTTAGAGAAGAAGATGGAAACAACATTGCTTTTC 183

Qy    181 TTTTCTCAAATAAATATGTGAATCAAAAGAAAAGAAAAACTTTTTTCAAGTTAATACATGGT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 TTTTCTCAAATAAATATGTGAATCAAAAGAAAAGAAAAACTTTTTTCAAGTTAATACATGGT 243

Qy    241 TCAGGAAAAGAAGAAACAGCAAAGAAGCCAAATCAGAGCTAAGGAAAAAAAAGAAAATAGA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 TCAGGAAAAGAAGAAACAGCAAAGAAGCCAAATCAGAGCTAAGGAAAAAAAAGAAAATAGA 303

Qy    301 CTTAAGTCTTCTTGTGTCAGAAACCTGAGTTTCATGAAGACACCCGCTCAGTAGATCTGGG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 CTTAAGTCTTCTTGTGTCAGAAACCTGAGTTTCATGAAGACACCCGCTCAGTAGATCTGGG 363

Qy    361 CACTTGGCCAAAGAAACAACAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGTGAATCATT 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 CACTTGGCCAAAGAAACAACAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGTGAATCATT 423

Qy    421 GACAAACTGCTTCCATAGATGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424 GACAAACTGCTTCCATAGATGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA 483

Qy    481 TTCAGTGAAGAAAATATTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 TTCAGTGAAGAAAATATTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA 543

Qy    541 CCTCAACAAATTCACCTTAAAGCAAAGCAATATATGAGAAATTATACGACTGATGCC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544 CCTCAACAAATTCACCTTAAAGCAAAGCAATATATGAGAAATTATACAGATGATGCC 603

Qy    601 CCAAAAGAGGTAACCTTGATTTTCACACAAAGGAAGTCATTACAAACAGCATCACTCAA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604 CCAAAAGAGGTAACCTTGATTTTCACACAAAGGAAGTCATTACAAACAGCATCACTCAA 663

Qy    661 CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664 CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC 723

Qy    721 AGTTATACACGTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAAGGAAGACCTCAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    724 AGTTATACACGTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAAGGAAGACCTCAG 783

Qy    781 AGACCAACAAATCTTAGGAGACGATCACGCTCATTCACCTGCAATGAATTCCAAGATGTA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    784 AGACCAACAAATCTTAGGAGACGATCACGCTCATTCACCTGCAATGAATTCCAAGATGTA 843

Qy    841 CAATCAGATGTTGCCATTTGGTTATAAAGAAAATGATTTTGCTCATTTTATGACAAAAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844 CAATCAGATGTTGCCATTTGGTTATAAAGAAAATGATTTTGCTCATTTTATGACAAAAC 903

Qy    901 TTTATACATTCTGCTTCTAACATATCGGCATGTTTATGTTAAGATTGGTCCATCCTTTAAA 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904 TTTATACATTCTGCTTCTAACATATCGGCATGTTTATGTTAAGATTGGTCCATCCTTTAA 963

Qy    961 CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAAACAAACTTTCTGCTAACAA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    964 CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAAACAAACTTTCTGCTAACAA 1023

Qy   1021 ATACACTACAGTATCTGCCAGTATATTTCTGTAAAACTTTCTTATTGATGTCATTCATTTA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1024 ATACACTACAGTATCTGCCAGTATATTTCTGTAAAACTTTCTTATTGATGTCATTCATTTA 1083

Qy   1081 TAATCAGAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAAGTAATAATGTTT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1084 TAATCAGAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAAGTAATAATGTTT 1143

Qy   1141 TATAAGATTGTAGAGTTAAGTAAGTTAAGTAAAAGTTAAAAGTTAAGTTGTCAAACA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1144 TATAAGATTGTAGAGTTAAGTAAGTTAAGTAAAAGTTAAAAGTTGTCAAACA 1203

Qy   1201 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATATATGTGTTATAAACATAATAAT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1204 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATATATGTGTTATAAACATAATAAT 1263

Qy   1261 ACTCAGATATCCAAATGTTCAGATAGCATTTTCAAGTTAATACATGT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1264 ACTCAGATATCCAAATGTTCAGATAGCATTTTCAAGTTAATACATGT 1323

Qy   1321 AATAGTTGTAGAAGTGATCTCGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1324 AATAGTTGTAGAAGTGATCTCGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT 1383
```

```
Qy  1381 ACTACTAACCCTGTCCCAAGAATAGTAATATCACCTCCAGTTATAAGGCAGCAACAGGAA 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1384 ACTACTAACCCTGTCCCAAGAATAGTAATATCACCTCCAGTTATAAGGCAGCAACAGGAA 1443

Qy  1441 CTTTTGTGAAGACACATTCATCTCTACAGAACTTCAGATTAAATATAATCTAGATTAATG 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1444 CTTTTGTGAAGACACATTCATCTCTACAGAACTTCAGATTAAATATAATCTAGATTAATG 1503

Qy  1501 ACTGAGAATAAGATCCACATTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1504 ACTGAGAATAAGATCCACATTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT 1563

Qy  1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTTGCATATCTCCAGGTAGGGA 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1564 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTTGCATATCTCCAGGTAGGGA 1623

Qy  1621 ACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGCCATCA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1624 ACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGCCATCA 1683

Qy  1681 ATGGCTTGAATAAAAACCAGAGAGAGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG 1740
         |||| ||||||||||||||  ||||||| ||||||||||||||||||||||||||||||
Db  1684 ATGGGCTGAATAAAAACCGAGA-GTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG 1742

Qy  1741 AATTGGGGTAGAAATCAGAAATGAGATGAGGGGAAGAAGCAAGGAGTCTAAGGCCCTAGC 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1743 AATTGGGGTAGAAATCAGAAATGAGATGAGGGGAAGAAGCAAGGAGTCTAAGGCCCTAGC 1802

Qy  1801 GATTTGGGCCATCGCCACATTGGTTCATATTCAGAAAG 1838
         ||||||||||||| ||||||||||||||||||||||||
Db  1803 GATTTGGGCCATTGCCACATTGGTTCATATTCAGAAAG 1840
```

RESULT 10
AL513175
LOCUS   AL513175   94175 bp   DNA   linear   PRI 15-NOV-2001
DEFINITION Human DNA sequence from clone RP11-142L4 on chromosome 1, complete
            sequence.
ACCESSION   AL513175
VERSION   AL513175.10 GI:16973028
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 94175)
AUTHORS   Martin,S.
TITLE   Direct Submission
JOURNAL   Submitted (15-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT   On Nov 16, 2001 this sequence version replaced gi:15620649.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e. phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-142L4 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-142L4 It may be shorter because we sequence overlapping
            sections only once, except for a short overlap.
            The true right end of clone RP11-142L4 is at 94175 in this
            sequence. The true left end of clone RP11-513B7 is at 2955 in this
            sequence. The true right end of clone RP11-198A7 is at 2000 in this
            sequence.
FEATURES   Location/Qualifiers
     source   1..94175
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-142L4"
                     /clone_lib="RPCI-11.1"
     misc_feature   28248..28327
                     /note="Sequence from overlapping clones RP11-513
                     (AL357566) and RP1-314M19 (AL391274). Assembly confirmed
                     by restriction digest."
     misc_feature   41580
                     /note="Tandem repeat. Forced join. Gap size estimated to
                     be approximately 50bp by restriction digest data."
ORIGIN

Query Match   71.3%; Score 1581; DB 9; Length 94175;
Best Local Similarity   99.1%; Pred. No. 4.9e-248;
Matches 1590; Conservative   0; Mismatches   15; Indels   0; Gaps   0;

Qy   608 AGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAACCTACCC 667
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2737 AGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAACCTACCC 2796

Qy   668 TCCACAGTTTGATGCTGCACAAAGCAGAGAGTGTATCAGCTCATGGAACAAGACAGTTATA 727
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2797 TCCACAGTTTGATGCTGCACAAAGCAGAGAGTGTATCAGCTCATGGAACAAGACAGTTATA 2856

Qy   728 CACGTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAAGGAAGACCTCGAGACCAA 787
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2857 CACGTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAAGGAAGACCTCGAGACCAA 2916

Qy   788 CAAATCTTAGGAGACGATCACGCTCATTACCTGCAATGAATTCCAAGATGTACAATCAG 847
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2917 CAAATCTTAGGAGACGATCACGCTCATTACCTGCAATGAATTCCAAGATGTACAATCAG 2976

Qy   848 ATGTTGCCATTGGTTATAAGAAAATTGATTTTGCTCATTTTATGACAAACTTATACA 907
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2977 ATGTTGCCATTGGTTATAAGAAAATTGATTTTGCTCATTTTATGACAAACTTATACA 3036

Qy   908 TCTGCTCTAACATATCGCATGTTATGTTAAGATTTGGTCCCATCCTTAAACTGAAAT 967
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3037 TCTGCTCTAACATATCGCATGTTATGTTAAGATTTGGTCCCATCCTTAAACTGAAAT 3096

Qy   968 ATGTCATGTGAAATTATTTAAAAATGTAAAACATAACAAAACTTCTGCTAACAAAATACATA 1027
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3097 ATGTCATGTGAAATTATTTAAAAATGTAAAACATAACAAAACTTCTGCTAACAAAATACATA 3156

Qy 1028 CAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTATGTCATTCCATTTATAATCAG 1087
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3157 CAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTATGATGTCATTCCATTTATAATCAG 3216

Qy 1088 AAAAAAACTTATTTCTAATCAAAGGCAGTACAAAAAAAGTAATAAATGTTTTATAAGA 1147
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3217 AAAAAAACTTATTTCTTAATCAAAGGCAGTACAAAAAAAGTACAAAAAAAGTAATAAATGTTTTATAAGA 3276

Qy 1148 TTGTAGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTGTCAAAAGTCT 1207
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 3277 TTGTAGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTGTCAAAAGTCT 3336
```

```
DEFINITION   Homo sapiens chromosome 1 clone RP1-314M19 map q31.1-31.3, 14
             unordered pieces.
ACCESSION    AL391274
VERSION      AL391274.4   GI:9931017
KEYWORDS     HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
AUTHORS      Mclay,K.
TITLE        Direct Submission
JOURNAL      Submitted (10-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
             CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
             requests: clonerequest@sanger.ac.uk
COMMENT      On Aug 27, 2000 this sequence version replaced gi:9926780.
             -------------- Genome Center
             Center: Sanger Centre
             Center code: SC
             Web site: http://www.sanger.ac.uk
             Contact: humquery@sanger.ac.uk
             -------------- Project Information
             Center project name: dJ314M19
             -------------- Summary Statistics
             Assembly program: XGAP4; version 4.5
             Sequencing vector: plasmid; L08752; 100% of reads Chemistry:
             Chemistry: Dye-terminator ET-amersham; 22% of reads Chemistry:
             Dye-terminator Big Dye; 77% of reads
             Consensus quality: 111500 bases at least Q40
             Consensus quality: 104399 bases at least Q30
             Consensus quality: 105813 bases at least Q20
             Insert size: 106619; sum-of-contigs
             Quality coverage: 3.51x in Q20 bases; sum-of-contigs
             --------------
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 14 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *        1      16334:  contig of 16334 bp in length
             *    16335      16434:  gap of 100 bp
             *    16435      23235:  contig of 6801 bp in length
             *    23236      23335:  gap of 100 bp
             *    23336      27979:  contig of 4644 bp in length
             *    27980      28079:  gap of 100 bp
             *    28080      45310:  contig of 17231 bp in length
             *    45311      45410:  gap of 100 bp
             *    45411      48030:  contig of 2620 bp in length
             *    48031      48130:  gap of 100 bp
             *    48131      52399:  contig of 4269 bp in length
             *    52400      52499:  gap of 100 bp
             *    52500      59052:  contig of 6553 bp in length
             *    59053      59152:  gap of 100 bp
             *    59153      66947:  contig of 7795 bp in length
             *    66948      67047:  gap of 100 bp
             *    67048      73293:  contig of 6246 bp in length
             *    73294      73393:  gap of 100 bp
             *    73394      79138:  contig of 5745 bp in length
             *    79139      79238:  gap of 100 bp
             *    79239      93778:  contig of 14540 bp in length
             *    93779      93878:  gap of 100 bp
             *    93879      97536:  contig of 3658 bp in length
             *    97537      97636:  gap of 100 bp
             *    97637     102981:  contig of 5345 bp in length
             *   102982     103081:  gap of 100 bp
             *   103082     107919:  contig of 4838 bp in length.
FEATURES             Location/Qualifiers
     source          1..107919
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
```

```
Qy   1208 AGTTGGGATTTTTACCAAAGCAGCATATATGTGTTATATAAACATAATAATACTTCAGA 1267
          ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   3337 AGTTGGGATTTTTACCAAAGCAGCATATATGTGTTATATAAACATAATAATAATACTTCAGA 3396

Qy   1268 TATCCAAATGTTCAGATAGCATTTTCATAATGAATGTCTCTTTTTTTTTCGTAATAGTG 1327
          ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   3397 TATCCAAATGTTCAGATAGCATTTTCATAATGAATGTCTCTTTTTTTTTCGTAATAGTG 3456

Qy   1328 TAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTTACTACTA 1387
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3457 TAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTTACTACTA 3516

Qy   1388 ACCCTGTCCAAGAATAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAACTTTTGT 1447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3517 ACCCTGTCCAAGAATAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAACTTTTGT 3576

Qy   1448 GAAGACACATTCATCTTCACAGAACTTCAGATTAAATAATCTAGATTAATGACTGAGA 1507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3577 GAAGACACATTCATCTTCACAGAACTTCAGATTAAATAATCTAGATTAATGACTGAGA 3636

Qy   1508 ATAAGATCCACATTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTATACAAAGT 1567
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3637 ATAAGATCCACATTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTATACAAAGT 3696

Qy   1568 ACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGTGGGGAACTAAGT 1627
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3697 ACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGTGGGGAACTAAGT 3756

Qy   1628 AGACTACCTTATCACCGGCTAGAAAAACTTGCTACTAAACATTAGGCCATCAATGGCTT 1687
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3757 AGACTACCTTATCACCGGCTAGAAAAACTTGCTACAAAACATTAGGCCATCAATGGCTT 3816

Qy   1688 GAATAAAAACCAGAGAAGGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAGAATTGGG 1747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3817 GAATAAAAACCAGAGAAGGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAGAATTGGG 3876

Qy   1748 GTAGAAATCAGAAATGAGATGAGGGGAAGAAGCAAGGAGTCTAAGCCCTAGCGATTTGG 1807
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3877 GTAGAAATCAGAAATGAGATGAGGGGAAGAAGCAAGGAGTCTAAGCCCTAGCGATTTGG 3936

Qy   1808 GCATCTCGCCACATTGGTTCATATTCAGAAAGTGTTATCCTCATTGATTATATTCTTGTTAA 1867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3937 GCATCTCGCCACATTGGTTCATATTCAGAAAGTGTTATCCTCATTGATTATATTCTTGTTAA 3996

Qy   1868 GCAAATCTCCTAAGTAATAATTATTATTCAAATAAGATTATACTTCATACATCTATATGTCAC 1927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3997 GCAAATCTCCTAAGTAATAATTATTATTCAAATAAGATTATACTTCATACATCTATATGTCAC 4056

Qy   1928 TGTTTTAAAGAGATATTTAATTTTAAATGTGTGTTACATGGTCTGTAAATATTTGTATTT 1987
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4057 TGTTTTAAAGAGATATTTAATTTTAAATGTGTGTTACATGGTCTGTAAATATTTGTATTT 4116

Qy   1988 AAAAATGCCATGCATTAGGCTTTGGAAATTTAATGTTAGTTGAAATGTAAAATGTGAAAA 2047
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4117 AAAAATGCCATGCATTAGGCTTTGGAAATTTAATGTTAGTTGAAATGTAAAATGTGAAAA 4176

Qy   2048 CTTTAGATCATTGTAGTAAGCTCTGACTGAAATATATTTTTAACTTCATTCATACAGTTATCTG 2107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4177 CTTTAGATCATTGTAGTAAGCTCTGACTGAAATATATTTTTAACTTCATTCATACAGTTATCTG 4236

Qy   2108 ACAATAAAAGCTCTGACAGTTGAATGTTGATTATCTTCCTATAATGTAATAAGGAATAAACA 2167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4237 ACAATAAAAGCTCTGACAGTTGAATGTTGATTATCTTCCTATAATGTAATAAGGAATAAACA 4296

Qy   2168 TTTTCTTCTTTTAGAGTAATAATGGGATAATAGAAAGTAGAAGTATTAATTGAG 2212
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   4297 TTTTCTTCTTTTAGAGTAATAATGGGATAATAGAAAGTAGAAGTATTAATTGAG 4341
```

RESULT 11
AL391274
LOCUS   AL391274   107919 bp   DNA   linear   HTG 11-SEP-2001

```
                    /db_xref="taxon:3606"
                    /chromosome="1"
                    /map="q31.1-31.3"
                    /clone="RP1-314M19"
                    /clone_lib="RPCI-1"
     misc_feature   1..16334
                    /note="assembly_fragment:00951
                    clone_end:T7
                    vector_side:left"
                    16435..23235
                    /note="assembly_fragment:00204"
     misc_feature   23336..27979
                    /note="assembly_fragment:00209"
     misc_feature   28080..45310
                    /note="assembly_fragment:00689"
     misc_feature   45411..48030
                    /note="assembly_fragment:00783"
     misc_feature   48131..52399
                    /note="assembly_fragment:00899"
     misc_feature   52500..59052
                    /note="assembly_fragment:00917"
     misc_feature   59153..66947
                    /note="assembly_fragment:00941"
     misc_feature   67048..73293
                    /note="assembly_fragment:01024"
     misc_feature   73394..79138
                    /note="assembly_fragment:01059"
     misc_feature   79239..93778
                    /note="assembly_fragment:01071"
     misc_feature   93879..97536
                    /note="assembly_fragment:01128"
     misc_feature   97637..102981
                    /note="assembly_fragment:01172"
     misc_feature   103082..107919
                    /note="assembly_fragment:01328"

ORIGIN

Query Match       71.3%;  Score 1581;  DB 2;  Length 107919;
Best Local Similarity 99.1%;  Pred. No. 4.8e-248;
Matches 1590;  Conservative 0;  Mismatches 15;  Indels 0;  Gaps 0;

Qy      608 AGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAACTACCC  667
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    42689 AGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAACTACCC 42748

Qy      668 TCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCGACGAGTGTATCGGAACAAGAGACAGTTATA  727
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    42749 TCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCGACGAGTGTATCGGAACAAGAGACAGTTATA 42808

Qy      728 CACGTTTTCTGAAATCTGACATCTATTAGACCTGATGGAAGGAGACCTCAGAGACCAA  787
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    42809 CACGTTTTCTGAAATCTGACATCTATTAGACCTGATGGAAGGAGACCTCAGAGACCAA 42868

Qy      788 CAAATCTTAGGAGACGATCACCGCTCATTTACCTGCAATGAATTCCAAGATGTACAATCAG  847
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    42869 CAAATCTTAGGAGACGATCACCGCTCATTTACCTGCAATGAATTCCAAGATGTACAATCAG 42928

Qy      848 ATGTTGCCATTGGTTATAAAGAAAATTGATTTGCTCATTTTATGACAAACTATACA  907
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    42929 ATGTTGCCATTGGTTATAAAGAAAATTGATTTGCTCATTTTATGACAAACTATACA 42988

Qy      908 TCTGCTTCTAACATATCCGCATGTTTATGTTAAGAATTGTAAGAAAAACCATCCTTTAAACTGAAAT  967
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    42989 TCTGCTTCTAACATATCCGCATGTTTATGTTAAGAATTGTAAGAAAAACCATCCTTTAAACTGAAAT 43048

Qy      968 ATGTCATGTGAAATTATTTTAAAATGTAAAACAAAACTTTCTGCTAACAAGAAAATACATA  1027
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43049 ATGTCATGTGAAATTATTTTAAAATGTAAAACAAAACTTTCTGCTAACAAGAAAATACATA 43108

Qy     1028 CAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTATAATCAG  1087
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43109 CAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTATAATCAG 43168

Qy     1088 AAAAAAAACTTATTTCTTAATCAAAGGCAGTACAAAAAAAAGTAATAATGTTTTATAAGA  1147
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43169 AAAAAAAACTTATTTCTTAATCAAAGGCAGTACAAAAAAAAGTAATAATGTTTTATAAGA 43228

Qy     1148 TTGTAGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTTGTCAAAGTTGTCAAACAAAGTCT  1207
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43229 TTGTAGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTTGTCAAAGTTGTCAAACAAAGTCT 43288

Qy     1208 AGTTGGGATTTTTACCAAAGCAGCATAATGTGTATATAAACATAATAATACTCAGA  1267
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43289 AGTTGGGATTTTTACCAAAGCAGCATAATGTGTATATAAACATAATAATACTCAGA 43348

Qy     1268 TATCCAAATGTTCAGATAGCATTTTCATAATGGGAGATCGAAGAACATTTATTATTGGGTACTACTA  1327
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43349 TATCCAAATGTTCAGATAGCATTTTCATAATGGGAGATCGAAGAACATTTATTATTGGGTACTACTA 43408

Qy     1328 TAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTACTACTA  1387
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43409 TAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTACTACTA 43468

Qy     1388 ACCCTGTCCCAGAAATAGTAATATCACCTCAGTTATAAGCCAGCAACAGGAACTTTTGT  1447
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43469 ACCCTGTCCCAGAAATAGTAATATCACCTCAGTTATAAGCCAGCAACAGGAACTTTTGT 43528

Qy     1448 GAAGACAATTCATCTCTACAGAACTTCAGATTAAATAATCTAGATTAATATAATGACTGAGA  1507
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43529 GAAGACAATTCATCTCTACAGAACTTCAGATTAAATAATCTAGATTAATATAATGACTGAGA 43588

Qy     1508 ATAAGATCCACATTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTATACAAAGT  1567
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43589 ATAAGATCCACATTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTATACAAAGT 43648

Qy     1568 ACTTCTGTTGGTCACAGAAACAAAACCAGATTTTGCATATCTCCAGTAGGGAACTAAGT  1627
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43649 ACTTCTGTTGGTCACAGAAACAAAACCAGATTTTGCATATCTCCAGTAGGGAACTAAGT 43708

Qy     1628 AGACTACCTTATCACCGGCTAAGAGAAACTTGCTACTAAACATTAGGCCATCAATGGCTT  1687
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43709 AGACTACCTTATCACCGGCTAAGAGAAACTTGCTACTAAACATTAGGCCATCAATGGCTT 43768

Qy     1688 GAATAAAAACCAGAGAGGTTTTCCCAGGACGTCCATGTTGTTGGCCCTTGGCCCTTAGAATTGGG  1747
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43769 GAATAAAAACCAGAGAGGTTTTCCCAGGACGTCCATGTTGTTGGCCCTTGGCCCTTAGAATTGGG 43828

Qy     1748 GTAGAAATCAGAAATGAGATGAGAGGGGAAGAACCAAGGAGTCTAAGGCCCTAGCCGATTTGG  1807
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43829 GTAGAAATCAGAAATGAGATGAGAGGGGAAGAACCAAGGAGTCTAAGGCCCTAGCCGATTTGG 43888

Qy     1808 GCATCTGCCACATTGGTTCATATTCAGAAAAGTGTTATCTCATTGATTATTCTTGTTAA  1867
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43889 GCATCTGCCACATTGGTTCATATTCAGAAAAGTGTTATCTCATTGATTATTCTTGTTAA 43948

Qy     1868 GCAAATCTCCTTAAGTAATTATTATTCAAATAAGATTCAAATAAGATTACTCATATATGTCAC  1927
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    43949 GCAAATCTCCTTAAGTAATTATTATTCAAATAAGATTCAAATAAGATTACTCATATATGTCAC 44008

Qy     1928 TGTTTAAAGAGATATTTAATTTTAAGTGTGTTACATGGTCTGTAAATATTTGTATTT  1987
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    44009 TGTTTAAAGAGATATTTAATTTTAAGTGTGTTACATGGTCTGTAAATATTTGTATTT 44068

Qy     1988 AAAAATGCCATTAGGCTTGGAAATTTAATGTTAGTTGAAATTTAAATGTAAAGTGAAAA  2047
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    44069 AAAAATGCCATTAGGCTTGGAAATTTAATGTTAGTTGAAATTTAAATGTAAAGTGAAAA 44128

Qy     2048 CTTTAGATCATTGTAGTAATAAATATTTTCGTCTAACAAGAAAATACATA  2107
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    44129 CTTTAGATCATTGTAGTAATAAATATTTTCGTCTAACAAGTTATCTG 44188

Qy     2108 ACAATAAAAGCTCTGACTGAATGTTGATTATCCTCCTATTATGTAATAAGGAATAAACA  2167
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    44189 ACAATAAAAGCTCTGACTGAATGTTGATTATCCTCCTATTATGTAATAAGGAATAAACA 44248

Qy     2168 TTTTCTTCTTTTAGTAAAAAAAAAAAAAAAAAAAAAAAAAGGGCG 2212
```

```
Db    44249 TTTTCTTCTTTTAGAGTAAGTAATATTGGGATAGTAATGAAAGAAGTATTAATTGAG 44293
            |||||||||||||||||         |  | ||| |  |   |||
```

RESULT 12
AX299963
LOCUS   AX299963   1486 bp   DNA   linear   PAT 26-NOV-2001
DEFINITION   Sequence 6 from Patent WO0183514.
ACCESSION   AX299963
VERSION   AX299963.1 GI:17129454
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Murray,D.L. and Gagnon.A.W.
  TITLE   Nucleic acids encoding a novel regulator of G protein signaling,
             rgs18, and uses thereof
  JOURNAL   Patent: WO 0183514-A 6 08-NOV-2001;
             Aventis Pharmaceuticals Products Inc. (US)
FEATURES   Location/Qualifiers
     source   1..1486
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   65.3%;   Score 1448.2; DB 6; Length 1486;
Best Local Similarity 99.4%;   Pred. No. 4.7e-226;
Matches 1464; Conservative   0; Mismatches   8; Indels   8; Gaps   1;

```
Qy    656 CTCAACTACCTCCCACAGTTTGATGCTGCACAAGCAGAGTGTATCAGCTCATCGAAC 715
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CTCAACTACCTCCCACAGTTTGATGCTGCACAAGCAGAGTGTATCAGCTCATCGAAC 60

Qy    716 AAGACAGTTATACACGTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAAGAC 775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAGACAGTTATACACGTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAAGAC 120

Qy    776 CTCGAGAACCAACAAATCTTAGGAGACGACATCACGCTCATTTACCTGCAATGAATTCCAAG 835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTCGAGAACCAACAAATCTTAGGAGACGACATCACGCTCATTTACCTGCAATGAATTCCAAG 180

Qy    836 ATGTACAATCAGATGTTGCCATTTGGTTATAAGAAAATTGATTTTGCTCATTTTTATGA 895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATGTACAATCAGATGTTGCCATTTGGTTATAAGAAAATTGATTTTGCTCATTTTTATGA 240

Qy    896 CAAACTTATACATCGCTTCTAACATATCGCATGTTTATGTTAAGATTTGGTTCCCATCCT 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAAACTTATACATCGCTTCTAACATATCGCATGTTTATGTTAAGATTTGGTTCCCATCCT 300

Qy    956 TTAAACTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACTTTCTGCTA 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTAAACTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACTTTCTGCTA 360

Qy   1016 ACAAAATACATACAGTATCTGCCAGTATATTCTAATCAAAAGGCAGTACAAAAAGTAATAA 1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ACAAAATACATACAGTATCTGCCAGTATATTCTAATCAAAAGGCAGTACAAAAAGTAATAA 420

Qy   1076 ATTTATAATCAGAGAAAAAAACTTATTTCTTAATCAAAAGTTAAATCAAAAAAAGTAATAA 1135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATTTATAATCAGAGAAAAAAACTTATTTCTTAATCAAAAGTTAAATCAAAAAAAGTAATAA 480

Qy   1136 TGTTTTATAAGATTGTGAGAGTTGTAGAGTTAAGTAAAAGTTAAGCTTTTGCAAAGTTC 1195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGTTTTATAAGATTGTGAGAGTTGTAGAGTTAAGTAAAAGTTAAGCTTTTGCAAAGTTC 540

Qy   1196 AAACAAAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAAATGTGTTATATAAACATA 1255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAACAAAAGTCTAGTTGGGATTTTTTACCAAAGCAGCATAAATGTGTTATATAAACATA 600

Qy   1256 ATAATACTTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTT 1315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    601 ATAATACTTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTT 660

Qy   1316 TTGGTAATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATT 1375
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTGGTAATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATT 720

Qy   1376 GGGTACTACTAACCCTGTCCAAGAAATAGTAATATCACCTCTAGTTATAAGCCAGCAAC 1435
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGGTACTACTAACCCTGTCCAAGAAATAGTAATATCACCTCTAGTTATAAGCCAGCAAC 780

Qy   1436 AGGAACTTTGTGAAGACACATTCATCTCTACAGAACTTCAGATTAAATATAATCTAGAT 1495
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGGAACTTTGTGAAGACACATTCATCTCTACAGAACTTCAGATTAAATATAATCTAGAT 840

Qy   1496 TAATGACTGAGAATAAGATCCACATTTGAACTCATTCCTAAGTGAACATGGACGTACCCA 1555
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TAATGACTGAGAATAAGATCCACATTTGAACTCATTCCTAAGTGAACATGGACGTACCCA 900

Qy   1556 GTTATACAAAGTACTTCTGTTGGTCACAGAAACATGACCCAGATTTGCATATCTCCAGT 1615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GTTATACAAAGTACTTCTGTTGGTCACAGAAACATGACCCAGATTTGCATATCTCCAGGT 960

Qy   1616 AGGGAACTAAGTAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGC 1675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AGGGAACTTTGTGAAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACATTAGGC 1020

Qy   1676 CATCAATGGCTGAATAAAAACCAGAGAGAGGTTTTCCCAGGAGCGTCTCATGTTTGGCCC 1735
          ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   1021 CATCAATGGCTGGAATAAAAACCAGAA-GTTTTTCCCAGGAGCGTCTCATGTTTGGCCC 1079

Qy   1736 TTTAGAATTGGGGTAGAAATCAGAAATGAAATCAGAAATGAGATGAGGGGAAGAGCAAGGAGTCTAAGGCC 1795
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1080 TTTAGAATTGGGGTAGAAATCAGAAATGAGATGAGGGGAAGAGCAAGGAGTCTAAGGCC 1139

Qy   1796 CTAGCGATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGATTA 1855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 CTAGCGATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGATTA 1199

Qy   1856 TATTCTTGTTGATGCAAATCTCCTTAAGTAATTATTATTAAATAGATTATACTCATACA 1915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200 TATTCTTGTTGATGCAAATCTCCTTAAGTAATTATTATTAAATAGATTATACTCATACA 1259

Qy   1916 TCTATATGTCACTGTTTTAAGAGATATTTAATTTTAATTTTTAATGTGTTTACATGGTCTGTAA 1975
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1260 TCTATATGTCACTGTTTTAAAGAGATATTTAATTTTTAATTTTTAATGTGTTTACATGGTCTGTAA 1319

Qy   1976 ATATTTGTATTTAAAAATGCCATGCATTAGGCTTTGGAAATTTAATGTTAGTTGAAATGT 2035
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1320 ATATTTGTATTTAAAAATGCCATGCATTAGGCTTTGGAAATTTAATGTTAGTTGAAATGT 1379

Qy   2036 AAAAATGAAAACTTTAGATCATTTGTAGTAATAAATATTTTAACTTCATTCATACAGT 2095
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380 AAAAATGAAAACTTTAGATCATTTGTAGTAATAAATATTTTAACTTCATTCATACAAT 1439

Qy   2096 TAAGTTTATCTGACAATAAAAGCTCTGACTGAA 2128
          |||||||||||||||||||||||||||||||||
Db   1440 TAAGTTTATCTGACAATAAAAGCTCTGACTGAA 1472
```

RESULT 13
AR165091
LOCUS   AR165091   1164 bp   DNA   linear   PAT 17-OCT-2001
DEFINITION   Sequence 3 from patent US 6274362.
ACCESSION   AR165091
VERSION   AR165091.1 GI:16238488
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 1164)
  AUTHORS   Hodge,M.R. and Yowe,D,
  TITLE   RGS-containing molecules and uses thereof

```
Db   941 GATGATTCATTTT------TAAAATACACCATGCAAATACATATTAAATGTAAGAACTT  994
Qy  1002 AAAACTTCTGCTAACAAAATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTA 1061
Db   995 TTTATATTATACTACAAAATAATTCATCATCTATCTTCCGAAATATTTTATGAAAATCTATC 1054
Qy  1062 TTTGATGTCATTCCATTTATATC 1085
Db  1055 TGATATTCTATTCTAATAAAATTC 1078
```

RESULT 14
BD265699
LOCUS   BD265699   1164 bp   DNA   linear   PAT 17-JUL-2003
DEFINITION RGS-containing molecules and uses thereof.
ACCESSION   BD265699
VERSION   BD265699.1 GI:33075467
KEYWORDS   JP 2002535979-A/2.
SOURCE   Mus sp.
  ORGANISM Mus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1164)
  AUTHORS   Hodge,M.R. and Yowe,D.
  TITLE   RGS-containing molecules and uses thereof
  JOURNAL   Patent: JP 2002535979-A 2 29-OCT-2002;
            MILLENNIUM PHARMACEUTICALS INC
COMMENT   OS   Mus sp. (mouse)
            PN   JP 2002535979-A/2
            PD   29-OCT-2002
            PF   04-FEB-2000 JP 2000597306
            PR   04-FEB-1999 US 09/244314
            PI   MARTIN R HODGE,DAVID YOWE
            PC   C12N15/09,C07K14/47,C07K16/18,C12N5/10,C12P21/02,C12Q1/02, PC
            C12Q1/68,
            PC   G01N33/53,G01N33/53,G01N33/566,C12N15/00,C12N15/00,C12N5/00 CC
            RGS-containing molecules and uses thereof
FEATURES   Location/Qualifiers
FH   Key   Location/Qualifiers
FT   CDS   (134)..(841).
                     /organism="Mus sp."
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10095"
ORIGIN

Query Match   26.9%; Score 596; DB 6; Length 1164;
Best Local Similarity 75.8%; Pred. No. 2e-87;
Matches 791; Conservative 0; Mismatches 220; Indels 33; Gaps 3;

```

JOURNAL     Patent: US 6274362-A 3 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1..1164
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match   26.9%;  Score 596;  DB 6;  Length 1164;
Best Local Similarity 75.8%;  Pred. No. 2e-87;
Matches  791;  Conservative  0;  Mismatches  220;  Indels  33;  Gaps  3;

```
Qy   69 TTATTCTACTATGTATATGTATATGTGAATAGTATTAATAAATGAACTAGGGAAGGATGTAAT 128
Db   41 TTTATTCAGGATGTTTTCCTATGAAATAGCATTCATCTGTGGGAGGAAGGACTAAGG   100
Qy 129 AAATTAGACATCTCTTCAT--TTAGAGAGAGATGGAAACAACATTGCTTTTCTTTTTCT   186
Db 101 AAATCTGACATCTGTTGGTCACTGGGACAGAATATGCTCACTGGTTTTCTTTCTCT   160
Qy 187 CAAATAAATATGTGAATCAAAAGAAAAAAACTTTTTCAAGTAATAACATGTTCAGGA   246
Db 161 CAATTAAATATGTGAATCAAAAGAGAAAAACTTTTTCAAACTAATGCATGGGTCAGGG   220
Qy 247 AAAGAAGAAACAAGCAAATCAGAGCTAAGGAAAAAAATAGACTAAGT   306
Db 221 AAAGAAGAAACAAGCATCGAGGCCAAATCAGAGCGAAAGAAAAAGGAATAGACTAAGT   280
Qy 307 CTTCTTGTGCAGAAACCTGAGTTTCATGAAGACTCTCGGCACTTG   366
Db 281 CTTCTCCTACAGAGGGCCTGACTTCCATGGAGAGGAGATCTCGCCCTCTTG   340
Qy 367 GCCAAAGAAACAAGAGTCTCCCTGAGAGGAGTCGAGTGAATCATTTGACAAA   426
Db 341 GCCAAAGAAACAAGAGTCTCTCCTGAGAAGAAGCAGTGAATCATTTGACAAA   400
Qy 427 CTGCTTCCCATAGAGATGGAATTTGGATAGCCTGTTTTACCAGATTCTTAAAACTGAATTCAGT 486
Db 401 TTGCTCTCCATAGAGATGGAGTGGATGCTTCTAAAACTGAATTCAGT   460
Qy 487 GAAGAAAATATTGAATTTGGATAGCCTGTGAGAAGCAAGGGACCTCAA   546
Db 461 GAGGAGAACATTGAATTTGGGTCGCCGTGTGAAGACTTCAAGAAGAACCTCAA   520
Qy 547 CAAATTCACCTTAAAGACAAAGCAATAATGAGAAATTATACAGACTGATGCCCCAAAA   606
Db 521 CAAATCATCCTAAAAGCAAAGGCAATCAATCCATGAGAATCATTCAGAAGAATGATGCCCCCAAA 580
Qy 607 GAGGTTAACCTGATTTTCACACAAAGAAGTCATTACAAACAGCATCACTCAACCTACC   666
Db 581 GAGGTTAACATTGATTTTTCATATCAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACT 640
Qy 667 CTCCACAGTTTGATGCTGCACAAAGACAGAGTGTATCAGCTCATGGAACAAGACAGTTAT 726
Db 641 CTCCACAGTTTGATACGGCACAAAGACAGAGTGTACCAGCTCATGGAACATGACAGTTAT 700
Qy 727 ACACGTTTTCTGAAATCTGACATCTTATTTAGACTTGATGGAAGGAACCTCAGAGACCA 786
Db 701 AAACGCTTTTTGAAATCTGAGACCTCACTTGATAGAGGAAGAGAGACCCTCAGAGACCA 760
Qy 787 ACAAATCTTAGGACGCATCACGCTCATTACCTGCAATGAAGATCCAGATGTACAATCA   846
Db 761 ACAAACCTTAGGACGACGATCACGATCATTACTTACTTCCAGATGAATGCATGGAATGTCA 820
Qy 847 GATGTTGCCATTTGGTTATAAGAAAATTGATTTTGCTCATTTTATGACAAACTTATAC   906
Db 821 GATGTTGCCATTTGGTTATGAGTAAAGTCATTTGTCTTTTTTGTATGTGTATGTGTAT   880
Qy 907 ATC-------------TGCTTCTAACATATCGCATGTTTATGTTAAGA   941
Db 881 ATCTAAAATATATACTAATACTAATGTGTACTTCTAAAAATATAGCTTGTGTGTATAAGAGA 940
Qy 942 TTTGGTCCATCCTTAAACTGAAAATGTCATGTGAAATTATTTTAAAAATGTAAAAAC 1001
```

```
Qy    69  TTATTCTACTATGTATATGTATGTATGAATAGTATTAATAAAATGAACTAGGAAGGATGTAAT  128
          ||||| ||||||| |||   |||||| ||||| |||||||||||||||||||||||||| ||
Db    41  TTTATTCAGGATGTTTCCATGAAATAGCTTTTCCATGAAATAGATTCATCTGTGGGAGAGGACTAAGG  100

Qy   129  AAATTAGACATCTCTTCAT--TTTAGAGAGGAAGATGGAAGCAACATTGCTTTCTTTTCT  186
          ||||| |||||||||||||  |||||||||| || ||| ||||||||||| ||||| |||
Db   101  AAATCTGACATCTGTTGGTCACTGGGACGAATATGGATATGTCACTGGTTTCTTTCTCT  160

Qy   187  CAAATAAATATGTGAATCAAAGAACTTTTTCAAGTAATACATGGTTCAGGA  246
          |||| |||||||||||||   || ||| ||||||| |||||||||| || 
Db   161  CAATTAAATATGTGAATCAATCAAAGAAACTTTTTCAAACTAATGCATGGGTCCAGGG  220

Qy   247  AAAGAGAAACAAGCAAAGAGAGCAAAATCAGAGCTAAGGAAAAAGAAATAGACTAAGT  306
          ||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||| |
Db   221  AAAGAGAAACAAGCATCGAGGCCAAAATCAGAGCCAAAGAAAAAGAAATAGAATAGACTAAGT  280

Qy   307  CTTCTTGTGCAGAAACCTGAGTTCATGAGCTCATGGAACTGAGAGATCGGGGACTTG  366
          ||||| |||||||||||||||||||| |||||||||||||||||||||||||||||| 
Db   281  CTTCTCCTACAGAGGCCTGACTTCCATGGAGCACCCGCTCCAGTGAGATCGGGCCCTCTG  340

Qy   367  GCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGTGAATTCATTTGACAAA  426
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   341  GCCAAAGAAACAAGAGTCTCTCCTGAAGAGCAGTGAAATGGGCTGAATTCATTTGACAAA  400

Qy   427  CTGCTTTCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCCAAGATGTACAATCA  486
          ||||| |||||||| |||| ||||||||||||| ||||||||||||||| |||||||
Db   401  TTGCTCTCCATAGAGATGGAGTGGATGTCTTCCGCCTGTGAAGATCTCAAGATGTAAAGTCA  460

Qy   487  GAAGAAATATTGGTTATAAAGCAAGGACCTCAA  546
          |||| ||||||||||||||||| ||||||||||
Db   461  GAGGAGAACATTGGTCGCCTGTGAAGACTTCAAGAGGAACCTCAA  520

Qy   547  CAAATTCACCTAAAGCAATCACGCTCATTTACCTGCAATGAGTGCCCCAAAA  606
          |||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   521  CAAATCATCCTAAAAGCAAAGGCAATCTATCAAAATGAGAAATTCATCAGAAATGAGTGCCCCCAAA  580

Qy   607  GAGGTAACCTGATTTGTGCATTACCTGCAATGAGTCATTACAAACAGCAGCACTCACCTACC  666
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   581  GAGGTTAACCTGATTTGTGCATTACTTGCAATGAGTAATTGCTAAGACCATCCGCCCAGCCCACT  640

Qy   667  CTCCACAGTTTGATGCTGCACAAAGAGTCTCCCCTGAAGAGGACCTCAGAGACCA  726
          |||||||| ||| |||||||||||||||||||||||| |||||||||||||||||
Db   641  CTCCACAGTTTGATACGGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTAT  700

Qy   727  ACACGTTTTCTGAAATCTGACAATCGCAAGATGTTTATGTTTAAGA  786
          ||||||||| ||| ||||| |||| |  || ||||| |||||||||
Db   701  AAAACGCTTTTGAAATCTGAGACATCTATTAGACTTGAGAAGGAAGACCTCAGAGACCA  760

Qy   787  ACAAATCTTAGGAGCAGATCACGCTCATTCACCTGCAATCAACTA  846
          ||||||||||||||||||||||||||||||||||||||||||||
Db   761  ACAAACCTTAGGAGGACGACGATCACGCGCATCGATCATTCAATGATTTCCAAGATGTACAATCA  820

Qy   847  GATGTTGCCATTTGGTTATAAAGCAAATTGATTTGCTCATTTTTATGACAAACTTATAC  906
          |||||||||||||||| |||   |||||||||||||||  || || |||||| || |||
Db   821  GATGTTGCCATTTGGTTATGAGTAAAAGTCATTTGCTTTCTTCTTTGATAGTGTATGTGTAT  880

Qy   907  ATC---------------------------TGCTTCTAACATATCGCATGTTTATGTTAAGA  941
          |||                           |||| || |||||| |||||||||||||||||
Db   881  ATCTAAAATATATACTAATACTAATGTGTACTTCTAAAATATAGCTTTGTGTATAAGAGAGA  940

Qy   942  TTTGGTCCCATCCTTTAAACTGAAATATGTGAAAATATTTTAAAAATGTAAAAAC  1001
          |||||| ||| || ||   |||||||||||||||||||| |||||||||||||
Db   941  GATGTTGCCATTTGCCATTTGGTTATGAGTAAAAGTCATTTCTTTTTTGATAGTGTATGTGTAT  994

Qy  1002  AAAACTTTCTGCTAACAAAATACACAGTATCTGCCAGTATATTCTGTAAAACCTTCTA  1061
          |  | |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   995  TTTATTATTACTAACTAAAATAATTCATCATCATCTTCCGAAATATTTTATGAAAATCTATC  1054

Qy  1062  TTTGATGTCATTCCATTTATAATC  1085
          |||||||||||||||||||||||
Db  1055  TGATATTCTATTCTAATAAAATTC  1078
```

RESULT 15
AR214936
LOCUS   AR214936   1164 bp   DNA   linear   PAT 25-SEP-2002
DEFINITION Sequence 3 from patent US 6410240.
ACCESSION   AR214936
VERSION   AR214936.1 GI:23312889
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1164)
  AUTHORS   Hodge,M.R. and Yowe,D.
  TITLE   RGS-containing molecules and uses thereof
  JOURNAL   Patent: US 6410240-A 3 25-JUN-2002;
FEATURES   Location/Qualifiers
     source   1..1164
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match   26.9%; Score 596; DB 6; Length 1164;
Best Local Similarity   75.8%; Pred. No. 2e-87;
Matches 791; Conservative   0; Mismatches 220; Indels 33; Gaps   3;

Mon Aug 23 07:18:08 2004

us-09-894-749-1.rge

Search completed: August 20, 2004, 09:58:55
Job time : 8791.76 secs

```
          GenCore version 5.1.6
      Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   August 20, 2004, 01:52:23 ; Search time 919.324 Seconds
                                       (without alignments)
                                       10244.763 Million cell updates/sec

Title:            US-09-894-749-1
Perfect score:    2217
Sequence:         1 gaattcggcttccatcctaa.........aaaaaaaaagggcggccgc 2217

Scoring table:    IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0

Searched:         3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:      6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq_29Jan04:*
            1:   geneseqn1980s:*
            2:   geneseqn1990s:*
            3:   geneseqn2000s:*
            4:   geneseqn2001as:*
            5:   geneseqn2001bs:*
            6:   geneseqn2002s:*
            7:   geneseqn2003as:*
            8:   geneseqn2003bs:*
            9:   geneseqn2003cs:*
           10:   geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2217	100.0	2217	3	AAA52089	Aaa52089 Human RGS
2	2217	100.0	2217	6	AAD42497	Aad42497 Human reg
3	2041.6	92.1	2144	6	AAS18340	Aas18340 Full leng
4	1816	81.9	1884	6	ABQ99387	Abq99387 Human cod
5	1753.2	79.1	1840	6	AAS18339	Aas18339 5'-RACE C
6	1448.2	65.3	1486	6	AAS18331	Aas18331 cDNA from
7	782	35.3	848	6	ABQ98617	Abq98617 Human ORF
8	596	26.9	1164	6	AAA52090	Aaa52090 Murine RG
9	596	26.9	1164	6	AAD42498	Aad42498 Murine re
10	227.2	10.2	241	6	AAS18334	Aas18334 cDNA enco
11	158.4	7.1	408	7	ACC43243	Acc43243 Nucleotid
12	158.4	7.1	1795	4	AAH76414	Aah76414 RGS prote
13	155.6	7.0	1345	6	ABT10881	Abt10881 Human CDN
14	155.6	7.0	1345	6	ABK83834	Abk83834 Human bre
15	155.6	7.0	1345	7	ACC46750	Acc46750 Human COP
16	155.6	7.0	1345	7	ACA56680	Aca56680 Signallin
17	155.6	7.0	1345	6	ADD14613	Add14613 Human src
18	155.6	7.0	1364	3	AAF16132	Aaf16132 Human pro
19	155.6	7.0	1382	9	ADB47313	Adb47313 Human cDN
20	153.2	6.9	745	2	AAV38084	Aav38084 Human reg
21	153.2	6.9	1691	3	AAZ36910	Aaz36910 cDNA enco
22	153.2	6.9	1923	2	AAX51745	Aax51745 DNA encod
23	153.2	6.9	1923	6	ABQ92597	Abq92597 Human sec
24	153.2	6.9	2075	7	ABX74396	Abx74396 Human cDN
25	152.4	6.9	1629	7	ABT42293	Abt42293 Toxicity
26	145.2	6.5	543	6	ABL99975	Abl99975 Rat disea
27	145.2	6.5	817	6	ABL99979	Abl99979 Rat disea
28	143.8	6.5	800	5	ACA56493	Aca56493 Human sig
29	143.8	6.5	1238	4	AAL26766	Aal26766 Human bre
30	143.8	6.5	2934	6	ABK47336	Abk47336 cDNA enco
31	143.8	6.5	3165	9	ABX63113	Abx63113 Human cDN
32	143.8	6.5	3167	9	ADD12660	Add12660 Human cDN
33	143.4	6.5	681	9	ADD34052	Add34052 Mouse mit
34	142.8	6.4	750	7	ACC43242	Acc43242 Nucleotid
35	141.2	6.4	1227	7	ABX74448	Abx74448 Human cDN
36	141.2	6.4	1329	9	ADE25703	Ade25703 Human cDN
37	141.2	6.4	1380	5	AAC78056	Aac78056 Human can
38	141.2	6.4	1398	6	AAF68404	Aaf68404 Human lun
39	141.2	6.4	1398	6	ABK38315	Abk38315 cDNA enco
40	141.2	6.4	1398	7	ABK64672	Abk64672 Human ben
41	141.2	6.4	1398	7	ACC46754	Acc46754 Human COP
42	141.2	6.4	1398	7	ACA56692	Aca56692 Signallin
43	141.2	6.4	1398	7	ACA10644	Aca10644 Human lun
44	141.2	6.4	1398	7	ABX99595	Abx99595 Lung canc
45	141.2	6.4	1398	10	ADE72378	Ade72378 Human lun

ALIGNMENTS

```
RESULT 1
AAA52089
AAA52089
ID   AAA52089 standard; cDNA; 2217 BP.
XX
AC   AAA52089;
XX
DT   04-DEC-2000  (first entry)
XX
DE   Human RGS protein coding sequence.
XX
KW   RGS; regulators of G-protein signaling; GTPase activating protein; GAP;
KW   G-alpha protein; cell adhesion; chemotaxis; vulnerary; immunosuppressor;
KW   anti-rheumatic; anti-arthritic; anti-diabetic; anti-inflamatory;
KW   cytostatic; hepatotropic; anti-anaemic; modulator; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH                     Location/Qualifiers
FT   Key
FT   CDS             160..867
FT                   /*tag= a
FT                   /product= "RGS_protein"
XX
PN   WO200046236-A2.
XX
PD   10-AUG-2000.
XX
PF   04-FEB-2000;  2000WO-US002977.
XX
PR   04-FEB-1999;  99US-00244314.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Hodge MR, Yowe D;
XX
DR   WPI; 2000-532893/48.
DR   P-PSDB; AAY97153.
XX
PT   Novel regulator of G-protein signaling nucleic acids and polypeptides,
PT   useful as diagnostic and investigative tools and to treat G-protein
PT   signaling disorders.
XX
PS   Claim 1; Page 100-102; 105pp; English.
XX
CC   The RGS (regulators of G-protein signaling) protein genes, clones
CC   AAH16395 and m1975, were identified in human and murine spleen cDNA
CC   libraries, respectively. Both proteins have unique N- and C-terminal
```

CC   sequences. The C-terminal location of the RGS domain is consistent with
CC   RGSs known to act as GTPase activating proteins (GAPs) for G-alpha
CC   proteins. G-alpha-i linked receptors support rapid adhesion and directed
CC   migration of leukocytes and other cell types. The novel RGS proteins may
CC   be used to modulate cell adhesion and chemotaxis, e.g. for aiding wound
CC   repair. The RGS proteins, related cDNAs and anti-RGS antibodies are
CC   useful for modulation, diagnosis and treatment of immune and respiratory
XX   disorders
SQ   Sequence 2217 BP; 801 A; 359 C; 380 G; 677 T; 0 U; 0 Other;

Query Match   100.0%; Score 2217; DB 3; Length 2217;
Best Local Similarity   100.0%; Pred. NO. 0;
Matches 2217; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 GAATTCGGCTTCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GAATTCGGCTTCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT 60

Qy   61 ATAACTTTTATTCTACTATGTATATGTATGGAATAGTATTAATAAATGAACTAGGGGAAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ATAACTTTTATTCTACTATGTATATGTATGGAATAGTATTAATAAATGAACTAGGGGAAG 120

Qy   121 GATGTAATAAATTAGACATCTCTTCATTTGAGAGAGAGATGGAAGCAACATTGCTTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GATGTAATAAATTAGACATCTCTTCATTTGAGAGAGAGATGGAAGCAACATTGCTTTTC 180

Qy   181 TTTTCTCAAATAAATATGTGTGAATCAAAAGAAAAAACTTTTTCAAGTTAATACATGGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TTTTCTCAAATAAATATGTGTGAATCAAAAGAAAAAACTTTTTCAAGTTAATACATGGT 240

Qy   241 TCAGGAAAGAAGAACAAGAAAGAGCCAAATCAGAGCCAAAATCAGAAGAAAAAGAAATAGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TCAGGAAAGAAGAACAAGAAAGAGCCAAATCAGAGCCAAAATCAGAAGAAAAAGAAATAGA 300

Qy   301 CTAAGTCTTCTTGTGCAGAAAACCTGAGTTTCATGAAGACACCCGTCCAGTAGATCTGGG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTAAGTCTTCTTGTGCAGAAAACCTGAGTTTCATGAAGACACCCGTCCAGTAGATCTGGG 360

Qy   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGAGGCAGTAGAATGGGGTGAAATCATTT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGAGGCAGTAGAATGGGGTGAAATCATTT 420

Qy   421 GACAAACTGCTTTCCCATAGAGATGGACTAGAGCCTTTTAAACTGAA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GACAAACTGCTTTCCCATAGAGATGGACTAGAGCCTTTTAAACTGAA 480

Qy   481 TTCAGTGAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAGCAAGGGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TTCAGTGAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAGCAAGGGA 540

Qy   541 CCTCAACAAATTCACCTTAAAGCAAAGCAAATTTATACAAACAGAAAATTATACAGACTGATGCC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCTCAACAAATTCACCTTAAAGCAAAGCAAATTTATACAAACAGAAAATTATACAGACTGATGCC 600

Qy   601 CCAAAAGAGGTAACCTTGATTTTCACCACAAAGAAGTCATTACAAACAGCATCACTCAA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CCAAAAGAGGTAACCTTGATTTTCACCACAAAGAAGTCATTACAAACAGCATCACTCAA 660

Qy   661 CCTACCCTCCACAGTTTGATGCTGCACAAAGCGACGTGTCATCAGCTTCAACAAGAC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CCTACCCTCCACAGTTTGATGCTGCACAAAGCGACGTGTCATCAGCTTCAACAAGAC 720

Qy   721 AGTTATACACGTTTTCTGAAATCTTAGGACAGACGATCATTAGACTTGATGGAGGAGACCTCAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AGTTATACACGTTTTCTGAAATCTTAGGACAGACGATCATTAGACTTGATGGAGGAGACCTCAG 780

Qy   781 AGACCAACAAATCTTAGGAGAGACGATCACGCTCATTAGACTTACCTGCAATGAATTCCAAGATGTA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGACCAACAAATCTTAGGAGAGACGATCACGCTCATTAGACTTACCTGCAATGAATTCCAAGATGTA 840

Qy   841 CAATCAGATGTTGCCATTGGTTATAAAGAAAAATTGATTTTGCTCATTTTTATGACAAAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CAATCAGATGTTGCCATTGGTTATAAAGAAAAATTGATTTTGCTCATTTTTATGACAAAC 900

Qy   901 TTTATACATCTGCTTCAACATATCGCATGTTGATGTTAAGATTGGTCCCATCCTTTAAA 960
          . |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTTATACATCTGCTTCAACATATCGCATGTTGATGTTAAGATTGGTCCCATCCTTTAAA 960

Qy   961 CTGAAATATGTCATGTGAAATATATTTAAAATATTTTAAAAATGTAAAAACAAAACTTTCTGCTAACAAA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CTGAAATATGTCATGTGAAATATATTTAAAATATTTTAAAAATGTAAAAACAAAACTTTCTGCTAACAAA 1020

Qy   1021 ATACATACAGTATCTGCCAGTATATTCTGTAAAACTTCGTAAAACTTCATTTGATGTCATTCATTTA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ATACATACAGTATCTGCCAGTATATTCTGTAAAACTTCGTAAAACTTCATTTGATGTCATTCATTTA 1080

Qy   1081 TAATCAGAAAAAAACTATTTCTTAATCAAAAGGCAGTACAAAAAAAGTAATAATGTTT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TAATCAGAAAAAAACTATTTCTTAATCAAAAGGCAGTACAAAAAAAGTAATAATGTTT 1140

Qy   1141 TATAAGATTGTAGAGTTAAGTAAAGTTAAGCTTTTGCAAAGTTGTCAAAGTTCAAACA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TATAAGATTGTAGAGTTAAGTAAAGTTAAGCTTTTGCAAAGTTGTCAAAGTTCAAACA 1200

Qy   1201 AAAGTCAGTTGGGATTTTTACCAAAGCAGCATATATGTGTATTATAAACATAATAAT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AAAGTCAGTTGGGATTTTTACCAAAGCAGCATATATGTGTATTATAAACATAATAAT 1260

Qy   1261 ACTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTTTGGT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ACTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTTTGGT 1320

Qy   1321 AATAGTGTAGAAGTGATCTGGTTCTACAATGGGAGATGAAGAACATTATTATTATTGGGTT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AATAGTGTAGAAGTGATCTGGTTCTACAATGGGAGATGAAGAACATTATTATTATTGGGTT 1380

Qy   1381 ACTACTAACCCTGTCCCAAGATAGTAATATCACCCTAGTTATAAGCCAGCAACAGGAA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ACTACTAACCCTGTCCCAAGATAGTAATATCACCCTAGTTATAAGCCAGCAACAGGAA 1440

Qy   1441 CTTTTGTGAAGACACATTCATCCTCTACAGAACTTCAGATTAAATAATAATCTAGATTAATG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CTTTTGTGAAGACACATTCATCCTCTACAGAACTTCAGATTAAATAATAATCTAGATTAATG 1500

Qy   1501 ACTGAGAATAAGATCCACATTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ACTGAGAATAAGATCCACATTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT 1560

Qy   1561 ACAAAGTACTTCTGTTGGTCACAGAAACATCACCAGATTTGCATATCTCCAGGTAGGGA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 ACAAAGTACTTCTGTTGGTCACAGAAACATCACCAGATTTGCATATCTCCAGGTAGGGA 1620

Qy   1621 ACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAAACATTAGGCCATCA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAAACATTAGGCCATCA 1680

Qy   1681 ATGGCTTGAATAAAAACCAGAGAAGGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 ATGGCTTGAATAAAAACCAGAGAAGGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG 1740

Qy   1741 AATTGGGTAGAAATCAGAAATCAGAAATGAGAGGGGAAGAAGCAAGGAGTCTAAGGCCCTAGC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 AATTGGGTAGAAATCAGAAATCAGAAATGAGAGGGGAAGAAGCAAGGAGTCTAAGGCCCTAGC 1800

Qy   1801 GATTTGGGCATCTGCCACATCTGGTTCATATTCAGAAGTGTATCTCATTGATTATATTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GATTTGGGCATCTGCCACATCTGGTTCATATTCAGAAGTGTATCTCATTGATTATATTC 1860

Qy   1861 TTGTTAAGCAAATCTCCTTAAGTAATTATTATTATTCAAATAAGATTATACTCATACATCTAT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TTGTTAAGCAAATCTCCTTAAGTAATTATTATTATTCAAATAAGATTATACTCATACATCTAT 1920

Qy   1921 ATGTCACTGTTTTAAAGAGATATTTAATTTTTAATGTGTGTTACATGGTCTGTAAATATT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ATGTCACTGTTTTAAAGAGATATTTAATTTTTAATGTGTGTTACATGGTCTGTAAATATT 1980

```
CC   useful for modulating the activity of RGS proteins and thus useful for
CC   treating immune and inflammatory disorders (Grave's disease, allergy,
CC   arthritis), respiratory disorders (asthma, pneumonia, sinusitis),
CC   haematological disorders (haematopoiesis, migration), platelet-associated
CC   disorders (thrombocytopaenia), invasive disorders (leukaemia),
CC   erythrocyte-associated disorders (anaemia), pancreatitis, hepatitis etc.
XX   The present sequence is human RGS protein, h16395 cDNA
SQ   Sequence 2217 BP; 801 A; 359 C; 380 G; 677 T; 0 U; 0 Other;

Query Match    100.0%;   Score 2217;   DB 6;   Length 2217;
Best Local Similarity  100.0%;   Pred. No. 0;
Matches 2217;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

Qy     1 GAATTCGGCTTCCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT    60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GAATTCGGCTTCCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT    60

Qy    61 ATAACTTTTATTCTACTATGTATATGTCGAATAGTATTAATAAATGAACTAGGGAAG   120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ATAACTTTTATTCTACTATGTATATGTCGAATAGTATTAATAAATGAACTAGGGAAG   120

Qy   121 GATGTAATAAATTAGACATCCTCCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GATGTAATAAATTAGACATCCTCCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC   180

Qy   181 TTTTCTCAAATAAATATGTGCAATCAAAGAGAAAAAACTTTTTCAAGTTAATACATGGT   240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TTTTCTCAAATAAATATGTGCAATCAAAGAGAAAAAACTTTTTCAAGTTAATACATGGT   240

Qy   241 TCAGGAAAGAGAAACAAGCAAAGAGCCAAATCAGAGCTAAGGAAAAAAGAAATAGA   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TCAGGAAAGAGAAACAAGCAAAGAGCCAAATCAGAGCTAAGGAAAAAAGAAATAGA   300

Qy   301 CTAAGTCTTCTGTGCAGAAACCTGAGTTCATGAAGACACCCGCTCCAGTAGATCTGGG   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTAAGTCTTCTGTGCAGAAACCTGAGTTCATGAAGACACCCGCTCCAGTAGATCTGGG   360

Qy   361 CACTTGGCCAAAGAACAAGAGTCTCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT   420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACTTGGCCAAAGAACAAGAGTCTCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT   420

Qy   421 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA   480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA   480

Qy   481 TTCAGTGAAGAAATATTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAGCAAGGGA   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TTCAGTGAAGAAATATTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAGCAAGGGA   540

Qy   541 CCTCAACAAATTCACCTTAAAGCAAAAGCAATATGAGAAATTTATACAGACTGATGCC   600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCTCAACAAATTCACCTTAAAGCAAAAGCAATATGAGAAATTTATACAGACTGATGCC   600

Qy   601 CCAAAAGAGGTTAACCTTGATTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGAACAAGAC   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CCAAAAGAGGTTAACCTTGATTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGAACAAGAC   660

Qy   661 CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGACAACAGAC   720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGACAACAGAC   720

Qy   721 AGTTATACACGTTTTCTGAAATCTGACATCTATTTAGACTTGAGTGGAGGAAGCAAGGACCTCAG   780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AGTTATACACGTTTTCTGAAATCTGACATCTATTTAGACTTGAGTGGAGGAAGCAAGGACCTCAG   780

Qy   781 AGACCAACAAATCTTAGGAGGAGACGGATCACGCTCATTACCTGCAATTCCAAGATGTA   840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGACCAACAAATCTTAGGAGGAGACGGATCACGCTCATTACCTGCAATTCCAAGATGTA   840

Qy   841 CAATCAGATGTTGCCATTTGGTTATAAAGAAAATTGATTTTGCTCATTTTATGACAAAC   900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CAATCAGATGTTGCCATTTGGTTATAAAGAAAATTGATTTTGCTCATTTTATGACAAAC   900
```

```
Qy  1981 TGTATTTAAAAATGCCATGCATTAGCCTTGGAAAATTAATGTTAGTTGAAATGTAAAAT  2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 TGTATTTAAAAATGCCATGCATTAGCCTTGGAAAATTAATGTTAGTTGAAATGTAAAAT  2040

Qy  2041 GTGAAAACTTAGATCATTGTGAGTAATAAAATATTTTAACTTCATTCATACAGTTAAGT  2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 GTGAAAACTTAGATCATTGTGAGTAATAAAATATTTTAACTTCATTCATACAGTTAAGT  2100

Qy  2101 TTATCTGCAATAAAAGCTCGACTGAATGTTGATTATCCTCCTATTATGTAATAAGGA  2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 TTATCTGCAATAAAAGCTCGACTGAATGTTGATTATCCTCCTATTATGTAATAAGGA  2160

Qy  2161 ATAAACATTTTCTTCTTCTTTTTAGAGTAAAAAAAAAAAAAAAAAAAAAAAGGCGGCCGC  2217
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 ATAAACATTTTCTTCTTCTTTTTAGAGTAAAAAAAAAAAAAAAAAAAAAAAGGCGGCCGC  2217
```

```
RESULT 2
AAD42497
ID   AAD42497 standard; cDNA; 2217 BP.
XX
AC   AAD42497;
XX
DT   15-NOV-2002 (first entry)
XX
DE   Human regulator of G-protein signalling (RGS) protein, h16395 cDNA.
XX
KW   Human; screening; RGS; regulator of G-protein signalling; pancreatitis;
KW   inflammatory disorder; allergy; Grave's disease; arthritis; sinusitis;
KW   respiratory disorder; asthma; pneumonia; therapy; immune disorder;
KW   haematological disorder; haematopoiesis; platelet-associated disorder;
KW   thrombocytopaenia; migration; invasive disorder; leukaemia; anaemia;
KW   erythrocyte-associated disorder; hepatitis; h16395; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             160..867
FT                   /tag= a
FT                   /product= "Human RGS protein"
XX
FN   US6410240-B1.
XX
PD   25-JUN-2002.
XX
PF   04-FEB-2000; 2000US-00498959.
XX
PR   04-FEB-1999; 99US-00244314.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Hodge MR, Yowe D;
XX
DR   WPI; 2002-588886/63.
DR   P-PSDB; AAE25829.
XX
PT   Screening assays for identifying agent that binds to human or mouse RGS
PT   (regulators of G-protein signalling) protein or its variant, or RGS
PT   protein encoded by homologous DNA sequences, or to host cell expressing
PT   the proteins.
XX
PS   Claim 15; Col 51-56; 42pp; English.
XX
CC   The invention relates to screening assays for identifying an agent that
CC   binds to: human or mouse RGS (regulators of G-protein signalling) protein
CC   or its variant, or RGS protein encoded by homologous DNA sequences; or a
CC   host cell expressing the RGS protein or its variant, or RGS protein
CC   encoded by homologous DNA sequences. The invention is useful for
CC   identifying an agent that binds to human or mouse RGS protein or its
CC   variant, or RGS protein encoded by homologous DNA sequences, or a host
CC   cell expressing the RGS protein or its variant, or RGS protein encoded by
CC   homologous DNA sequences. The agents identified using the invention are
```

```
Qy   1981 TGTATTTAAAAATGCCATGCATTAGGCTTTGAAATTAAGTGTTAGTGAAATGTAAAAT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TGTATTTAAAAATGCCATGCATTAGGCTTTGAAATTAAGTGTTAGTGAAATGTAAAAT 2040

Qy   2041 GTGAAAACTTTAGATCATTGTAGTAATAATAATATTTTAACTTCATTCATACAGTTAAGT 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GTGAAAACTTTAGATCATTGTAGTAATAATAATATTTTAACTTCATTCATACAGTTAAGT 2100

Qy   2101 TTATCTGACAATAAAAAGCTCTGACTGAAGTTGATTATCCTTCCTATTATGTAATAAGGA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 TTATCTGACAATAAAAAGCTCTGACTGAAGTTGATTATCCTTCCTATTATGTAATAAGGA 2160

Qy   2161 ATAAACATTTCTTCTTTTAGAGTAAAAAAAAAAAAAAAAAAAAAAGGGGCGGCCGC 2217
          |||||||||||||||||||||||||||||||
Db   2161 ATAAACATTTCTTCTTTTAGAGTAAAAAAAAAAAAAAAAAAAAAAGGGGCGGCCGC 2217


RESULT 3
AAS18340
ID   AAS18340 standard; cDNA; 2144 BP.
XX
AC   AAS18340;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Full length cDNA encoding human RGS18.
XX
KW   Human; regulator of G protein signalling; RGS18; arterial thrombosis;
KW   platelet activation dysfunction; myocardial infarction; stroke;
KW   coronary artery disease; cerebrovascular disease; unstable angina;
KW   deep vein thrombosis; systemic thromboembolism; anti-coagulant;
KW   invasive cardiac procedure; ss.
XX
OS   Homo sapiens.
XX
FH                    Location/Qualifiers
FT   CDS             163..870
FT                   /tag= a
FT                   /product= "RGS18 polypeptide"
XX
PN   WO200183514-A2.
XX
PD   08-NOV-2001.
XX
PF   26-APR-2001; 2001WO-US013540.
XX
PR   28-APR-2000; 2000US-0207786P.
PR   02-AUG-2000; 2000GB-00018833.
XX
PA   (AVET ) AVENTIS PHARM PROD INC.
XX
PI   Murray DL, Gagnon AW;
XX
DR   WPI; 2002-055453/07.
DR   P-PSDB; AAU10749.
XX
PT   Isolated regulator of G protein signaling polypeptide, useful for
PT   prevention/treatment of platelet activation dysfunction, such as arterial
PT   thrombosis, myocardial infarction, coronary artery disease and stroke.
XX
PS   Claim 1; Fig 1; 127pp; English.
XX
CC   The present invention relates to the isolation of polynucleotide
CC   sequences, that encode a novel regulator of G protein signalling (RGS)
CC   polypeptide, RGS18, from human platelets. The invention also provides
CC   nucleotide primers and probes specific for an RGS18 nucleic acid. The
CC   sequences of the invention are useful for the manufacture of a medicament
CC   for the prevention or treatment of a platelet activation dysfunction,
CC   such as arterial thrombosis, myocardial infarction, coronary artery
CC   disease, stroke, cerebrovascular disease, unstable angina, deep vein
CC   thrombosis, systemic thromboembolism, and in invasive cardiac procedures
CC   for anti-coagulant purposes. The sequences of the invention can also be
CC   used for the detection of RGS18 nucleic acids and RGS18 polypeptides. The
```

```
Qy    901 TTATACATCGCTTCTCAACTATCCATGTTATGTTAAGATTGGTCCCATCCTTTAAA 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTATACATCGCTTCTCAACTATCCATGTTATGTTAAGATTGGTCCCATCCTTTAAA 960

Qy    961 CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACTTTCTGCTAACAA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACTTTCTGCTAACAA 1020

Qy   1021 ATACATACAGTAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTGATGTCATTCCATTA 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ATACATACAGTAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTGATGTCATTCCATTA 1080

Qy   1081 TAATCAGAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAGTAATAAATGTTT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TAATCAGAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAGTAATAAATGTTT 1140

Qy   1141 TATAAGATTGTGAGTTAAGTAAAGTTAAGCTTTTGCAAAGTTGTCAAAGTTCAAACA 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TATAAGATTGTGAGTTAAGTAAAGTTAAGCTTTTGCAAAGTTGTCAAAGTTCAAACA 1200

Qy   1201 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAATATGTTATATAAAACATAATAAT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAATATGTTATATAAAACATAATAAT 1260

Qy   1261 ACTTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTTTTTGGT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ACTTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTTTTTGGT 1320

Qy   1321 AATAGTGTAGAAGTGATCTGGTTCTTACAAATGGGAGATGAAGAACATTTATTATTGGGTT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AATAGTGTAGAAGTGATCTGGTTCTTACAAATGGGAGATGAAGAACATTTATTATTGGGTT 1380

Qy   1381 ACTACTAACCCTGTCCCAAGAAGTAATATCACCCTCAGTTATAAGCCAGCAACAGGAA 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ACTACTAACCCTGTCCCAAGAAGTAATATCACCCTCAGTTATAAGCCAGCAACAGGAA 1440

Qy   1441 CTTTTGTTGAAGACACATTCATCTCTACAGAACTTCAGATTAAATATAATTCTAGATTAATG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CTTTTGTTGAAGACACATTCATCTCTACAGAACTTCAGATTAAATATAATTCTAGATTAATG 1500

Qy   1501 ACTGAGAATAAGATCCACCATTGGTTGAACTCATTCATTGAACTGAACATGAACGTACCCAGTTAT 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ACTGAGAATAAGATCCACCATTGGTTGAACTCATTCATTGAACTGAACATGAACGTACCCAGTTAT 1560

Qy   1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGAAGATTTGCATATCTCCAGGTAGGGA 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGAAGATTTGCATATCTCCAGGTAGGGA 1620

Qy   1621 ACTAAGTAGACTACCCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGCCATCA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ACTAAGTAGACTACCCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGCCATCA 1680

Qy   1681 ATGGCTTGAATAAAACCAGAGAGAGGTTTTCCCAGACGTCTCATGTTTCGGCCCCTTAG 1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 ATGGCTTGAATAAAACCAGAGAGAGGTTTTCCCAGACGTCTCATGTTTCGGCCCCTTAG 1740

Qy   1741 AATTGGGGTAGAAATCAGAAATGAGATGAGGGGAAGAAGCAAGAGAGTCTAAGGCCCTAGC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 AATTGGGGTAGAAATCAGAAATGAGATGAGGGGAAGAAGCAAGAGAGTCTAAGGCCCTAGC 1800

Qy   1801 GATTTGGGCATCGCCACCATTGGTTCATATTGGTTCATTGATTATTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GATTTGGGCATCGCCACCATTGGTTCATATTGGTTCATTGATTATATTC 1860

Qy   1861 TTGTTAAGCAAATCTCCTAAGTAATTATTATTCAAATAAGATTACTCATACATCTAT 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TTGTTAAGCAAATCTCCTAAGTAATTATTATTCAAATAAGATTACTCATACATCTAT 1920

Qy   1921 ATGTCACTGTTTTAAAGAGATATTTAATTTTTAATGTGTGTTACATGGTCGTAAATATT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ATGTCACTGTTTTAAAGAGATATTTAATTTTTAATGTGTGTTACATGGTCGTAAATATT 1980
```

```
CC  present sequence represents the full length cDNA sequence encoding RGS18
CC  from human platelets
XX
SQ  Sequence 2144 BP; 784 A; 340 C; 365 G; 655 T; 0 U; 0 Other;

Query Match     92.1%;    Score 2041.6;    DB 6;    Length 2144;
Best Local Similarity   99.5%;    Pred. No. 0;
Matches 2058;   Conservative   0;   Mismatches   9;   Indels   1;   Gaps   1;
```

```
Qy    61  ATAACTTTTATTCTACTAATGTATATGTATGGAATAGTATTAATAAATGAACTAGGGAAG   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    64  ATAACTTTTATTCTACTAATGTATATGTATGGAATAGTATTAATAAATGAACTAGGGAAG   123

Qy   121  GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAAGAAGATGAAACAACATTGCTTTC   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   124  GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAAGAAGATGAAACAACATTGCTTTC   183

Qy   181  TTTTCTCAAATAAATATGTGAATCAAAGAGAAAAACTTTTTCAAGTTAATACATGGT      240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184  TTTTCTCAAATAAATATGTGAATCAAAGAGAAAAACTTTTTCAAGTTAATACATGGT      243

Qy   241  TCAGGAAAAGAAGAAGCAAAGAGCCAAAATCAGAGCTAAGGAAAAAAGAAAATAGA      300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   244  TCAGGAAAAGAAGAAGCAAAGAGCCAAAATCAGAGCTAAGGAAAAAAGAAAATAGA      303

Qy   301  CTAAGTCTTCTTCTGTGCAGAAACCTGAGTTCATGAGGAGACACCCGCTCCAGTAGATCTCGGG   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   304  CTAAGTCTTCTTCTGTGCAGAAACCTGAGTTCATGAGGAGACACCCGCTCCAGTAGATCTCGGG   363

Qy   361  CACTTGGCCAAAGAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364  CACTTGGGCAAAGAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT   423

Qy   421  GACAAACTGCTTTCCCATAGAATGGACTAGAGCGTTTACCAGATTTCTTAAAACTGAA    480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   424  GACAAACTGCTTTCCCATAGAATGGACTAGAGCGTTTACCAGATTTCTTAAAACTGAA    483

Qy   481  TTCAGTGAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGA    540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   484  TTCAGTGAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGA    543

Qy   541  CCTCAACAAATTCACCTTAAAGCAAATACAATATGAGAATTTGAAATTTATGAGAATTTATCAGACTGATGCC   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   544  CCTCAACAAATTCACCTTAAAGCAAATACAATATGAGAATTTGAAATTTATGAGAATTTATCAGACTGATGCC   603

Qy   601  CCAAAAGAGGTTAACCTTGATTTCTGCACAAAGAAGTCATTACAAAACAGCATCACTCAA   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   604  CCAAAAGAGGTTAACCTTGATTTCTGCACAAAGAAGTCATTACAAAACAGCATCACTCAA   663

Qy   661  CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAATATCTATTTGACTTACCTGGAACAGAGGAAGAACAAGAC   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   664  CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAATATCTATTTGACTTACCTGGAACAGAGGAAGAACAAGAC   723

Qy   721  AGTTATACACGTTTTCATCGCCATGTAGACTGACTTAGACTTTAGACTCTCATGATGGAAGGAAGACCTCAG   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   724  AGTTATACACGTTTTCATCGCCATGTAGACTGACTTAGACTTTAGACTCTCATGATGGAAGGAAGACCTCAG   783

Qy   781  AGACCAACAAACTTAGGAGACGATCACGCTCATTTACCTGCCAATGAATGAATTCCAAGATGTA   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   784  AGACCAACAAACTTAGGAGACGATCACGCTCATTTACCTGCCAATGAATGAATTCCAAGATGTA   843

Qy   841  CAATCAGATGTGCCATTTGCCATTCGCTCATTTTGCTCATTTTATGTCTCATTTTATGACAAAC   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   844  CAATCAGATGTGCCATTTGCCATTCGCTCATTTTGCTCATTTTATGTCTCATTTTATGACAAAC   903

Qy   901  TTATCAATCTCGCTTCGCTTCTCAACATATCGCCATGTTTGGTCCCATCCTTTAAA   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   904  TTATCAATCTCGCTTCGCTTCTCAACATATCGCCATGTTTGGTCCCATCCTTTAAA   963

Qy   961  CTGAAATATGTCATGTGAAATTATTTAAAAATGTAAAAAACTTTCTGCTAACAAA   1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   964  CTGAAATATGTCATGTGAAATTATTTAAAAATGTAAAAAACTTTCTGCTAACAAA   1023
```

```
Qy  1021  ATACACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCATTTA   1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1024  ATACACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCATTTA   1083

Qy  1081  TAATCAGAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAGTAATAATGTTT   1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1084  TAATCAGAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAGTAATAATGTTT   1143

Qy  1141  TATAAGATTGTAGAGTTAAGTAAAAGTTAAGTTTTGCAAAAGTTGTCAAAAGTTCAAACA   1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1144  TATAAGATTGTAGAGTTAAGTAAAAGTTAAGTTTTGCAAAAGTTGTCAAAAGTTCAAACA   1203

Qy  1201  AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATATATGTGTTATATAAACATAATAAT   1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1204  AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAATATGTGTTATATAAACATAATAAT   1263

Qy  1261  ACTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTTTTTTGGT   1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1264  ACTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTTTTTTGGT   1323

Qy  1321  AATAGTGTAGAGTGATCTGGTTCTTACAATGGGAGTGAAGAACATTTATTATTGGGGTT   1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1324  AATAGTGTAGAGTGATCTGGTTCTTACAATGGGAGTGAAGAACATTTATTATTGGGGTT   1383

Qy  1381  ACTACTAACCCTGTCCCAAGAATAGTAATATCACCTCTAGTATAAGCCAGCAACAGGAA   1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1384  ACTACTAACCCTCTCCCAAGAATAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAA   1443

Qy  1441  CTTTGTGAAGACACATTCATCCTCACGAGAACTTCAGATTAATATATCTAGATTAATG   1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1444  CTTTTGTGAAGACACATTCATCCTCACGAGAACTTCAGATTAATATATCTAGATTAATG   1503

Qy  1501  ACTGAGAATAAGATCCACATTGAACTCCTAAGTGAACATGACGTACCCAGTTAT   1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1504  ACTGAGAATAAGATCCACATTGAACTCCTAAGTGAACATGACGTACCCAGTTAT   1563

Qy  1561  ACAAAGTACTTCTGTTGGTCACAGAAACGATTTGCATATCTCCAGGTAGGAGGA   1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1564  ACAAAGTACTTCTGTTGGTCACAGAAACGATTTGCATATCTCCAGGTAGGAGGA   1623

Qy  1621  ACTAAGTAGACTACCTTATCACCGGCTAAGAAACTTGCTACTAAACTATTAGGCCATCA   1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1624  ACTAAGTAGACTACCTTATCACCGGCTAAGAAACTTGCTACTAAACTATTAGGCCATCA   1683

Qy  1681  ATGGCTGAATAAAACCAGAGAGGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG   1740
          ||||||||||||||||||||||||||||||       ||||||||||||||||||
Db  1684  ATGGGTGAATAAAACCAGAA-GTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG   1742

Qy  1741  AATTGGGGTAGAAATCAGAAATGAGAAATGAGAGGGGAAGAACAAGGAGTCTAAGGCCCTAGC   1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1743  AATTGGGGTAGAAATCAGAAATGAGAAATGAGAGGGGAAGAACAAGGAGTCTAAGGCCCTAGC   1802

Qy  1801  GATTTGGGCATCTGCCACATTGGTTCATATTGGTCCATCCTTTATTATCTCATTGATTATTC   1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1803  GATTTGGGCATCTGCCACATTGGTTCATATTGGTCCATCCTTTATTATCTCATTGATTATTC   1862

Qy  1861  TTGTTAAGCAAATCTCCTTAAAGAGATATTTAATTTTAATGTGTTACATGGTCTGTAAATATT   1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1863  TTGTTAAGCAAATCTCCTTAAAGAGATATTTAATTTTAATGTGTTACATGGTCTGTAAATATT   1922

Qy  1921  ATGTCACTGTTTAAAAATGCCATGCATTTGGAAATTTAATGTTAGTTGAAAATGTAAAAT   1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1923  ATGTCACTGTTTAAAAATGCCATGCATTTGGAAATTTAATGTTAGTTGAAAATGTAAAAT   1982

Qy  1981  TGTATTTAAAAAATGCCATGGCATTTGGTCCATCCTTTAAA   2040
          |||||||||||||||||||||||||||||||||||||||||
Db  1983  TGTATTTAAAAAATGCCATGGCATTTGGTCCATCCTTTAAA   2042

Qy  2041  GTGAAAACTTAGATCCATCATTTGTAGTAGTAATAAATATTTTAACTTCATTCAGTTAAGT   2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2043  GTGAAAACTTAGATCCATCATTTGTAGTAGTAATAAATATTTTAACTTCATTCACAATTAAGT   2102

Qy  2101  TTATCTGACAATAAAGCTCTGACTGAA   2128
```

Db   2103 TTATCTGACACATAAAAGCTCTGACTGAA 2130
            ||||||||||||||||||||||||||||||

RESULT 4
ABQ99387
ID   ABQ99387 standard; cDNA; 1884 BP.
XX
AC   ABQ99387;
XX
DT   25-FEB-2003 (first entry)
XX
DE   Human coding sequence SEQ ID 120.
XX
KW   Human; expressed sequence tag; EST; chromosome 16;
KW   haematopoietic disorder; central nervous system disease; viral infection;
KW   peripheral nervous system disease; non-healing wound; infectious disease;
KW   immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW   fungal infection; autoimmune disorder; coagulation disorder; nootropic;
KW   antiallergic; antinflammatory; immunosuppressive; neuroprotective;
KW   cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW   immunostimulant; cerebroprotective; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200259260-A2.
XX
PD   01-AUG-2002.
XX
PF   16-NOV-2001; 2001WO-US042950.
XX
PR   17-NOV-2000; 2000US-00714936.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI   Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR   WPI; 2002-590824/63.
DR   N-PSDB; ABP64801.
XX
PT   New isolated polynucleotide, useful in research, diagnostic or
PT   therapeutic methods, e.g. preventing or treating disorders involving
PT   aberrant protein expression or biological activity.
XX
PS   Claim 1; SEQ ID NO 120; 394pp; English.
XX
CC   The present invention relates to novel human coding sequences (ABQ99268-
CC   ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC   therapeutic, diagnostic and research methods. The polynucleotides may be
CC   used in the field of molecular biology as hybridisation probes, primers
CC   for PCR, for chromosome and gene mapping, for the recombinant production
CC   of protein, or in generation of anti-sense DNA or RNA. The
CC   polynucleotides are useful in diagnostics as expressed sequence tags
CC   (ESTs) for identifying expressed genes or for physical mapping of the
CC   human genome. The proteins may be used as molecular weight markers, or as
CC   nutritional sources or supplements. The proteins may be used to maintain
CC   and expand cell population in a totipotential or pluripotential state
CC   useful for re-engineering damaged or diseased tissues, transplantation,
CC   manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC   polynucleotides and proteins are useful for preventing, treating or
CC   ameliorating disorders involving aberrant protein expression or
CC   biological activity, e.g. haematopoietic disorders, central/peripheral
CC   nervous system diseases, mechanical and traumatic disorders, non-healing
CC   wounds, immune deficiencies and disorders, infectious diseases caused by
CC   viral, bacterial or fungal infection, autoimmune disorders, allergic
CC   reactions and conditions, coagulation disorders, or cancer. The
CC   polynucleotide sequences of the invention were assembled from ESTs
CC   isolated mainly by sequencing by hybridisation, and in some cases,
CC   sequences obtained from one or more public databases. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences

XX
SQ   Sequence 1884 BP; 681 A; 314 C; 328 G; 561 T; 0 U; 0 Other;

      Query Match   81.9%; Score 1816; DB 6; Length 1884;
      Best Local Similarity 99.9%; Pred. No. 0;
      Matches 1827; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy   61 ATAACTTTTTATTCTACTATGTATATGTATGGAATAGTATTAATAAATGACTAGGGAAG 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   58 ATAACTTTTTATTCTACTATGTATATGTATGGAATAGTATTAATAAATGACTAGGGAAG 117

Qy   121 GATGTAATAAATTAGACATCTCTCATTTTAGAGAGAGGATGGAAACAACATTGCTTTTC 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   118 GATGTAATAAATTAGACATCTCTCATTTTAGAGAGAGGATGGAAACAACATTGCTTTTC 177

Qy   181 TTTTTCAAATAAATATGTGTGAATCAAAGAAAGAAAAAACTTTTTCAAGTTAATACATGGT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   178 TTTTTCAAATAAATATGTGTGAATCAAAGAAAGAAAAAACTTTTTCAAGTTAATACATGGT 237

Qy   241 TCAGGAAAAGAAGAAACAAGCAAGAGAGCCAAAATCAGAGGTAAGAAAAAAGAAATAGA 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   238 TCAGGAAAAGAAGAAACAAGCAAGAGAGCCAAAATCAGAGGTAAGAAAAAAGAAATAGA 297

Qy   301 CTAAGTCTTCTGTGCAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGG 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   298 CTAAGTCTTCTGTGCAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGG 357

Qy   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAATGGGTGAATCATTT 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   358 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAATGGGTGAATCATTT 417

Qy   421 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   418 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA 477

Qy   481 TTCAGTGAAGAAAATATTGATTTTGGATAGCCTGTGAAGATTCAAGAAAAGCAAGGGA 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   478 TTCAGTGAAGAAAATATTGATTTTGGATAGCCTGTGAAGATTCAAGAAAAGCAAGGGA 537

Qy   541 CCTCAACAAATTCACCTTAAAGCAAAAGCAATATATGAGAAATTTATACAGACTGATGCC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   538 CCTCAACAAATTCACCTTAAAGCAAAAGCAATATATGAGAAATTTATACAGACTGATGCC 597

Qy   601 CCAAAAGAGGTTAACCTTGATTTTCCACACAAAAGAAGTCATTACAAAACAGCATCACTCAA 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   598 CCAAAAGAGGTTAACCTTGATTTTCCACACAAAAGAAGTCATTACAAAACAGCATCACTCAA 657

Qy   661 CCTACCCTCCACAGTTTGATGCTGCACAAGCAGCAGAGTGTATCAGCTCATGGAACAAGAC 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   658 CCTACCCTCCACAGTTTGATGCTGCACAAGCAGCAGAGTGTATCAGCTCATGGAACAAGAC 717

Qy   721 AGTTATACACGTTTCTGAAATCTGACATCTATTTGATGGAAGGAAGACCTCAG 780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   718 AGTTATACACGTTTCTGAAATCTGACATCTATTTGATGGAAGGAAGACCTCAG 777

Qy   781 AGACCAACAAATCTTAGGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   778 AGACCAACAAATCTTAGGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA 837

Qy   841 CAATCAGATGTGCCATTTGGTTATAAAGAAAATGATTTTGCTCATTTTTATGCACAAAC 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   838 CAATCAGATGTGCCATTTGGTTATAAAGAAAATGATTTTGCTCATTTTTATGCACAAAC 897

Qy   901 TTATACATCTGCTTCTAACATATGCCATGTTTATGTTAAGATTTGGTCCCCATCCTTTAA 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   898 TTATACATCTGCTTCTAACATATGCCATGTTTATGTTAAGATTTGGTCCCCATCCTTTAA 957

Qy   961 CTGAAATATGTCATGTGAAATTATTTAAAAATGTAAAAACAAAACTTTCTGCTAACAA 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   958 CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACTTTCTGCTAACAA 1017

Qy   1021 ATACATACAGTATCTGCCAGTATATTCGTAAAACCTTCATTCTATTGATGTCATTCCATTTA 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
KW   invasive cardiac procedure; 5'-RACE; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183514-A2.
XX
PD   08-NOV-2001.
XX
PF   26-APR-2001; 2001WO-US013540.
XX
PR   28-APR-2000; 2000US-0200786P.
PR   02-AUG-2000; 2000GB-0018833.
XX
PA   (AVET ) AVENTIS PHARM PROD INC.
XX
PI   Murray DL, Gagnon AW;
XX
DR   WPI; 2002-055453/07.
XX
PT   Isolated regulator of G protein signaling polypeptide, useful for
PT   prevention/treatment of platelet activation dysfunction, such as arterial
PT   thrombosis, myocardial infarction, coronary artery disease and stroke.
XX
PS   Claim 1; Page 116-117; 127pp; English.
XX
CC   The present invention relates to the isolation of polynucleotide
CC   sequences, that encode a novel regulator of G protein signalling (RGS)
CC   polypeptide, RGS18, from human platelets. The invention also provides
CC   nucleotide primers and probes specific for an RGS18 nucleic acid. The
CC   sequences of the invention are useful for the manufacture of a medicament
CC   for the prevention or treatment of a platelet activation dysfunction,
CC   such as arterial thrombosis, myocardial infarction, coronary artery
CC   disease, stroke, cerebrovascular disease, unstable angina, deep vein
CC   thrombosis, systemic thromboembolism, and in invasive cardiac procedures
CC   for anti-coagulant purposes. The sequences of the invention can also be
CC   used for the detection of RGS18 nucleic acids and RGS18 polypeptides. The
CC   present sequence represents 5'-RACE clone DNA used to generate the full
CC   length human RGS18 cDNA
XX
SQ   Sequence 1840 BP; 668 A; 309 C; 328 G; 535 T; 0 U; 0 Other;

Query Match             79.1%;  Score 1753.2;  DB 6;  Length 1840;
Best Local Similarity   99.5%;                 Pred. No. 8.8e-304;
Matches 1769;           Conservative    0;  Mismatches    8;  Indels    8;  Gaps    1;

Qy    61 ATAACTTTTATTCTACTATGTATATGTGAATAGTATTAATAAATGAACTAGGGAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    64 ATAACTTTTATTCTACTATGTATATGTGAATAGTATTAATAAATGAACTAGGGAAG 123

Qy   121 GATGTAATAAATTAGACATCTTCATTTAGAGAGAAGAATGGAAACAACATTGCTTTC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   124 GATGTAATAAATTAGACATCTTCATTTAGAGAGAAGAATGGAAACAACATTGCTTTC 183

Qy   181 TTTTCTCAAATAAATATGTGAATCAAAGAAGAGCAAAATCAGAGCTAAGGAAAAAGAAATAGA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184 TTTTCTCAAATAAATATGTGAATCAAAGAAGAGCAAAATCAGAGCTAAGGAAAAAGAAATAGA 243

Qy   241 TCAGGAAAAGAAGAAAGAGAAACAAGCAAGCAAGAGCCAAAATCAGAGCTAAGGAAAAAGAAATAGA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   244 TCAGGAAAAGAAGAAAGAGAAACAAGCAAGCAAGAGCCAAAATCAGAGCTAAGGAAAAAGAAATAGA 303

Qy   301 CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAAGATCTGG 360
         |||||||||||||||||||||||||||||||||||||||||||
Db   304 CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAAGATCTGG 363

Qy   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT 423

Qy   421 GACAAACTGCTTTCCCATAGAGATGGACTAGAGCTTTACCAGATTTCTAAAACTGAA 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   424 GACAAACTGCTTTCCCATAGAGATGGACTAGAGCTTTACCAGATTTCTAAAACTGAA 483
```

```
Db  1018 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTA 1077

Qy  1081 TAATCAGAAAAAAACTTATTTCTTAATCAAAGGCAGTACAAAAAAGTAATAAATGTTT 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1078 TAATCAGAAAAAAACTTATTTCTTAATCAAAGGCAGTACAAAAAAGTAATAAATGTTT 1137

Qy  1141 TATAAGATTGTAGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTTGTCAAAAGTTCAAACA 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1138 TATAAGATTGTAGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTTGTCAAAAGTTCAAACA 1197

Qy  1201 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAATATGTGTTATATAAACATAATAAT 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1198 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAATATGTGTTATATAAACATAATAAT 1257

Qy  1261 ACTCAGATATCCAAATGTTCAGATAGTCAGATGCATTTTCATATAGTCTCTTTTTTTTGGT 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1258 ACTCAGATATCCAAATGTTCAGATAGTCAGATGCATTTTCATATAGTCTCTTTTTTTTGGT 1317

Qy  1321 AATAGTGTAGAAGTGATCTGGTTCTTACAGATGGGAGATGAAGAACATTTATTATTGGGTT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1318 AATAGTGTAGAAGTGATCTGGTTCTTACAGATGGGAGATGAAGAACATTTATTATTGGGTT 1377

Qy  1381 ACTACTAACCCTGCCCAAGAATAGTAATATCACCCTCAGTTATAAGCCAGCAACAGGAA 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1378 ACTACTAACCCTGCCCAAGAATAGTAATATCACCCTCAGTTATAAGCCAGCAACAGGAA 1437

Qy  1441 CTTTTGTGAAGACACATCATCCTCACAGAACTCAGATTAAATAATCTAGATTAATG 1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1438 CTTTTGTGAAGACACATCATCCTCACAGAACTCAGATTAAATAATCTAGATTAATG 1497

Qy  1501 ACTGAGAATAAGATCCACATTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1498 ACTGAGAATAAGATCCACATTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT 1557

Qy  1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTTGCATATCTCCAGTAGGGA 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1558 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTTGCATATCTCCAGTAGGGA 1617

Qy  1621 ACTAAGTAGAGTACTTTATCACCGGCTAAGAGAAAACTTGCTACTAAACTATTAGGCCATCA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1618 ACTAAGTAGAGTACTTTATCACCGGCTAAGAGAAAACTTGCTACTAAACTATTAGGCCATCA 1677

Qy  1681 ATGGCTTGAATAAAAACCAGAGAGAGGTTTTCCCAGACGTCTCATGTTTGGCCCTTTAG 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1678 ATGGCTTGAATAAAAACCAGAGAGAGGTTTTCCCAGACGTCTCATGTTTGGCCCTTTAG 1737

Qy  1741 AATTGGGGTAGAAATCAGAAATCAGAGATGAGGGAAGAGAGCAAGGAGTCTAAGGCCCTAGC 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1738 AATTGGGGTAGAAATCAGAAATCAGAGATGAGGGAAGAGAGCAAGGAGTCTAAGGCCCTAGC 1797

Qy  1801 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGATTATATTC 1860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1798 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGATTATATTC 1857

Qy  1861 TTGTTAAGCAAATCTCCTTAAGTAATTA 1888
         ||||||||| ||||||||||||||||||
Db  1858 TTGTTAAGC-AATCTCCTTAAGTAATTA 1884
```

```
RESULT 5
AAS18339
ID   AAS18339 standard; DNA; 1840 BP.
XX
AC   AAS18339;
XX
DT   12-MAR-2002 (first entry)
XX
DE   5'-RACE clone DNA used to generate full length human RGS18 cDNA.
XX
KW   Human; regulator of G protein signalling; RGS18; arterial thrombosis;
KW   platelet activation dysfunction; myocardial infarction; stroke;
KW   coronary artery disease; cerebrovascular disease; unstable angina;
KW   deep vein thrombosis; systemic thromboembolism; anti-coagulant;
```

```
Qy    481  TTCAGTGAAGAAATATATGAATTTGGATAGCCTGTGTGAAGATTTCAAGAAAGCAAGGGA  540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484  TTCAGTGAAGAAATATATGAATTTGGATAGCCTGTGTGAAGATTTCAAGAAAGCAAGGGA  543

Qy    541  CCTCAACAAATTCACCTTAAAGCAAAGCAATATATGAGAAATTTATACAGACTGATGCC   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544  CCTCAACAAATTCACCTTAAAGCAAAGCAATATATGAGAAATTTATACAGACTGATGCC   603

Qy    601  CCAAAAGAGGTTAACCTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAA   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604  CCAAAAGAGGTTAACCTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAA   663

Qy    661  CCTACCCCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAAGGAACAAGAC  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664  CCTACCCCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGTGATGGAACAAGAC  723

Qy    721  AGTTATACACGTTTCTGAAATCTGACATCTATTTAGACTTAGACTTACCTGACAATGAATTCCAAGATGTA  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    724  AGTTATACACGTTTCTGAAATCTGACATCTATTTAGACTTAGACTTACCTGACAATGAATTCCAAGATGTA  783

Qy    781  AGACCAACAAATCTTAGGAGACGACGATCACGCTCATTACCTGACAATGAATTCCAAGATGTA  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    784  AGACCAACAAATCTTAGGAGACGACGATCACGCTCATTACCTGACAATGAATTCCAAGATGTA  843

Qy    841  CAATCAGATGTTGCCATTGGTTATAAAGAAAATTGATTTTGCTCATTTTATGACAAAC   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844  CAATCAGATGTTGCCATTGGTTATAAAGAAAATTGATTTTGCTCATTTTATGACAAAC   903

Qy    901  TTATACATCTGCTTCTAACATATCGCATGTTATGTTAAGATTTGGTCCCATCCTTTAAA   960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904  TTATACATCTGCTTCTAACATATCGCATGTTATGTTAAGATTTGGTCCCATCCTTTAAA   963

Qy    961  CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACTTTCTGCTAACAAA  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    964  CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACTTTCTGCTAACAAA  1023

Qy   1021  ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTA  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1024  ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTA  1083

Qy   1081  TAATCAGAAAAAAAACTTATTCTTAATCAAAGGCAGTACAAAAAAGTAATAATGTTT  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1084  TAATCAGAAAAAAAAACTTATTCTTAATCAAAGGCAGTACAAAAAAAGTAATAATGTTT  1143

Qy   1141  TATAAGATTGTAGAGTTAAGTAAAAAGTTAAGTAAGTAAGCGCAAAAGTGTCAAAAGTTCAAACA  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1144  TATAAGATTGTAGAGTTAAGTAAAAAGTTAAGTAAGTAAGCGCAAAAGTGTCAAAAGTTCAAACA  1203

Qy   1201  AAAGTCTAGTTGGGATTTTTACCAAGCAGCATAATATGTGTTATATAAACATAATAAT  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1204  AAAGTCTAGTTGGGATTTTTTACCAAGCAGCATAATATGTGTTATATAAACATAATAAT  1263

Qy   1261  ACTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCCTTTTTTTTTGGT  1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1264  ACTCAGATATCCAAATGTTCAGATAGCATTTTTCATAATGAATGTTCCTTTTTTTTTTGGT  1323

Qy   1321  AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1324  AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT  1383

Qy   1381  ACTACTAACCCTGTCCCAGAAGATAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAA  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1384  ACTACTAACCCTGTCCCAGAAGAATAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAA  1443

Qy   1441  CTTTTGTGAAGACACACATTCATCTCTACAGAACTTCAGATTAATATATAATCTAGATTAATG  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1444  CTTTTGTGAAGACACACATTCATCTCTACAGAACTTCAGATTAATATATAATCTAGATTAATG  1503

Qy   1501  ACTGAGAATAAGATCCACATTCCACATTGAACTCATTCCTAAGTGAACATGGAACGTACCCCAGTTAT  1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1504  ACTGAGAATAAGATCCACATTCCACATTGAACTCATTCCTAAGTGAACATGGAACGTACCCAGTTAT  1563

Qy   1561  ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGGTAGGGA  1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1564  ACAAAGTACTTCTGTTGGTCACAGAAACATCACCAGATTTGCATATCTAAACTATTAGGCCATCA  1623

Qy   1621  ACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGCCATCA  1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1624  ACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGCCATCA  1683

Qy   1681  ATGGCTTGAATAAAAACCAGAGAAGGTTTTCCCAGGAGTCTCATGTTTGGCCCTTTAG  1740
           |||||||||||||||||   ||||||||| ||||||||||||||||||||||||||||
Db   1684  ATGGGCTGGAATAAAAACCCGAGAA-GTTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG  1742

Qy   1741  AATTGGGTAGAAATCAGAAATGAGAAATGAGATGAGGGAGGAAGAAGCAAGAGTCAAGGCCTAGC  1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1743  AATTGGGTAGAAATCAGAAATGAGAAATGAGATGAGGGAGGAAGAAGCAAGAGTCAAGGCCCTAGC  1802

Qy   1801  GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAG  1838
           ||||||||||||||||||||||||||||||||||||||
Db   1803  GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAG  1840
```

RESULT 6
```
AAS18331
ID   AAS18331 standard; cDNA; 1486 BP.
XX
AC   AAS18331;
XX
DT   12-MAR-2002  (first entry)
XX
DE   cDNA from a human thyroid library encoding C-terminal region of RGS18.
XX
KW   Human; regulator of G protein signalling; RGS18; arterial thrombosis;
KW   platelet activation dysfunction; myocardial infarction; stroke;
KW   coronary artery disease; cerebrovascular disease; unstable angina;
KW   deep vein thrombosis; systemic thromboembolism; anti-coagulant;
KW   invasive cardiac procedure; thyroid; ss.
XX
OS   Homo sapiens.
XX
FH                       Location/Qualifiers
FT   Key
FT   CDS              3..212
FT                    /tag= a
FT                    /partial
FT                    /product= "Carboxy-terminal region of RGS18"
FT                    /note= "This sequence lacks a start codon"
XX
PN   WO200183514-A2.
XX
PD   08-NOV-2001.
XX
PF   26-APR-2001; 2001WO-US013540.
XX
PR   28-APR-2000; 2000US-0200786P).
PR   02-AUG-2000; 2000GB-00018833.
XX
PA   (AVET ) AVENTIS PHARM PROD INC.
XX
PI   Murray DL, Gagnon AW;
XX
DR   WPI; 2002-055453/07.
DR   P-PSDB; AAU10748.
XX
PT   Isolated regulator of G protein signaling polypeptide, useful for
PT   prevention/treatment of platelet activation dysfunction, such as arterial
PT   thrombosis, myocardial infarction, coronary artery disease and stroke.
XX
PS   Example 2; Page 113-114; 127pp; English.
XX
CC   The present invention relates to the isolation of polynucleotide
CC   sequences, that encode a novel regulator of G protein signalling (RGS)
CC   polypeptide, RGS18, from human platelets. The invention also provides
CC   nucleotide primers and probes specific for an RGS18 nucleic acid. The
CC   sequences of the invention are useful for the manufacture of a medicament
CC   for the prevention or treatment of a platelet activation dysfunction,
```

```
CC    such as arterial thrombosis, myocardial infarction, coronary artery
CC    disease, stroke, cerebrovascular disease, unstable angina, deep vein
CC    thrombosis, systemic thromboembolism, and in invasive cardiac procedures
CC    for anti-coagulant purposes. The sequences of the invention can also be
CC    used for the detection of RGS18 nucleic acids and polypeptides. The
CC    present sequence represents a cDNA sequence from an Incyte clone derived
CC    from a human thyroid library. The sequence is used to isolate the full
CC    length cDNA encoding RGS18 from human platelets
XX
SQ    Sequence 1486 BP; 529 A; 235 C; 240 G; 482 T; 0 U; 0 Other;

Query Match             65.3%;  Score 1448.2;  DB 6;  Length 1486;
Best Local Similarity   99.4%;  Pred. No. 2.4e-249;
Matches 1464;  Conservative   0;  Mismatches    8;  Indels    1;  Gaps    1;

Qy      656  CTCAACCTACCCTCCACAGTTTGATGCTGCACAAGCAGAGTGTATCAGCTCATGGAAC  715
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1  CTCAACCTACCCTCCACAGTTTGATGCTGCACAAGCAGAGTGTATCAGCTCATGGAAC   60

Qy      716  AAGACAGTTATACACGTTTTCTGAAATCGACATCTATTTAGACTTGATGGAAGGAAGAC  775
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61  AAGACAGTTATACACGTTTTCTGAAATCGACATCTATTTAGACTTGATGGAAGGAAGAC  120

Qy      776  CTCAGAGACCAACAAATCTTAGGAGACCAGCGATCACGCTCATTTACCTGCAATGAATTCCAAG  835
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121  CTCAGAGACCAACAAATCTTAGGAGACCAGCGATCACCCTCATTTACCTGCAATGAATTCCAAG  180

Qy      836  ATGTACAATCAGATGTTGCCATTGGTATAAAGAAAATTGATTTGCTCATTTTTATGA  895
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181  ATGTACAATCAGATGTTGCCATTGGTATAAAGAAAATTGATTTGCTCATTTTTATGA  240

Qy      896  CAAACTTATACATCTGCTTCTAACATATCGCATGTTATGTAAGATTGGTCCCATCCT  955
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241  CAAACTTATACATCTGCTTCTAACATATCGCATGTTATGTAAGATTGGTCCCATCCT  300

Qy      956  TTAAACTGAAAATGTCATGTGAAATTATTTTTAAAAATGTAAAAACAAAACTTTCTGCTA  1015
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301  TTAAACTGAAAATGTCATGTGAAATTATTTTTAAAAATGTAAAAACAAAACTTTCTGCTA  360

Qy     1016  ACAAAATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCC  1075
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361  ACAAAATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCC  420

Qy     1076  ATTTATAATCAGAAAAAAAAACTTATTCTTAATCAAAAGGCAGTACAAAAAAGTAATAA  1135
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421  ATTTATAATCAGAAAAAAAAACTTATTCTTAATCAAAAGGCAGTACAAAAAAGTAATAA  480

Qy     1136  TGTTTTATPAAGATTGTAGAGTTAAGTAAGTAAAAGTTAAGCTTTTGCAAAGTTCAAAAGTTC  1195
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481  TGTTTTATPAAGATTGTAGAGTTAAGTAAGTAAAAGTTAAGCTTTTGCAAAGTTCAAAAGTTC  540

Qy     1196  AAACAAAAAGTCTAGTTGGGGATTTTTACCAAAGCAGCATAAATATGTGTTATAAAACATA  1255
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541  AAACAAAAAGTCTAGTTGGGGATTTTTACCAAAGCAGCATAAATATGTGTTATAAAACATA  600

Qy     1256  ATAATACTTCAGATATCCAAATGTTCAGATAGCATTTTCATATAGAATGTTCTCTTTTTT  1315
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601  ATAATACTTCAGATATCCAAATGTTCAGATAGCATTTTCATATAGAATGTTCTCTTTTTT  660

Qy     1316  TTGGTAATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATT  1375
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661  TTGGTAATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATT  720

Qy     1376  GGGTTACTTACTAACCCTGTCCCAAGAATAGTAATATCACCTCCAGTTATAAGCCAGCAAC  1435
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721  GGGTTACTTACTAACCCTGTCCCAAGAATAGTAATATCACCTCCAGTTATAAGCCAGCAAC  780

Qy     1436  AGGAACTTTTGTGAAGACACATCTCTACAGAACTTCAGATAAATAATAATCTAGAT  1495
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781  AGGAACTTTTGTGAAGACACATCTCTACAGAACTTCAGATAAATAATAATCTAGAT  840

Qy     1496  TAATGACTGAGAATAAGATCCACATTTGAACTCATTCCTCAGTGAACTGGACGTACCCA  1555
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841  TAATGACTGAGAATAAGATCCACATTTGAACTCATTCCTCAGTGAACATGGACGTACCCA  900

Qy     1556  GTTATACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGT  1615
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901  GTTATACAAAGTACTTCGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGGT  960

Qy     1616  AGGGAACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGC  1675
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961  AGGGAACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGC  1020

Qy     1676  CATCAAATGGCTTGAATAAAAAACCAGAGAGAGGTTTTCCCAGGACGTCTCATGTTTGGGCCC  1735
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     1021  CATCAATGGGCCTGGAATAAAAAACCGAGAA-GTTTTCCCAGGACGTCTCATGTTTGGGCCC  1079

Qy     1736  TTTAGAATTGGGGTAGAAAATCAGAAATGAGATGAGGGGAAGGGGAAGCAAGGAGTCTAAGGCC  1795
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1080  TTTAGAATTGGGGTAGAAATCAGAAATGAGAAATGAGATGAGGGGAAGGGGAAGCAAGGAGTCTAAGGCC  1139

Qy     1796  CTAGCGATTGGCATCTGCCACATTGGTTCATATTCGAGAAAGTGTTATCTCATTGATTA  1855
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1140  CTAGCGATTTGGGCATCTGCCACATTGGTTCATATTCGAGAAAGTGTTATCTCATTGATTA  1199

Qy     1856  TATTCTTGTTAAGCAAATCTCCTTAAGTAATAATTATTATTCAAATAAGATTATACTCATACA  1915
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1200  TATTCTTGTTAAGCAAATCTCCTTAAGTAATAATTATTATTCAAATAAGATTATACTCATACA  1259

Qy     1916  TCTATATGTCACTGTTTTAAAGAGATATTTAATTTTTTAATGTGTGTTACATGGTCTGTAA  1975
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1260  TCTATATGTCACTGTTTTAAAGAGATATTTAATTTTTTAATGTGTGTTACATGGTCTGTAA  1319

Qy     1976  ATATTTGTATTTAAAAAATGCCATGCATTTGTTTTGAAATTAAATGCATTTAAATGTGAAATGT  2035
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1320  ATATTTGTATTTAAAAAATGCCATGCATTTGTTTTGAAATTAAATGCATTTAAATGTGAAATGT  1379

Qy     2036  AAAATGTGAAAACTTTAGTAGTCATTGTAGTAATAATAATATTTTTAACTTCATTCATTCATACAGT  2095
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1380  AAAATGTGAAAACTTTAGTAGTCATTGTAGTAATAATAATATTTTTAACTTCATTCATTCATACAAT  1439

Qy     2096  TAAGTTTATCTGACAATAAAAGCTCTGACTGAA  2128
             |||||||||||||||||||||||||||||||||
Db     1440  TAAGTTTATCTGACAATAAAAGCTCTGACTGAA  1472
```

```
RESULT 7
ABQ98617
ID   ABQ98617 standard; DNA; 848 BP.
XX
AC   ABQ98617;
XX
DT   04-NOV-2002  (first entry)
XX
DE   Human ORF424 coding sequence.
XX
KW   Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW   Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW   human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW   cancer; cardiovascular disease; allergy; autoimmune disease;
KW   wound healing; blood coagulation disorder; inflammatory disorder; ds.
XX
OS   Homo sapiens.
XX
PN   US2002082206-A1.
XX
PD   27-JUN-2002.
XX
PF   30-MAY-2001; 2001US-00867550.
XX
PR   30-MAY-2000; 2000US-0208427P.
XX
PA   (LEAC/) LEACH M D.
PA   (MEHR/) MEHRABAN F.
PA   (CONL/) CONLEY\P B.
PA   (TOPP/) TOPPER J N.
```

PA   (LAWD/) LAW D.
XX
PI   Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
DR   WPI; 2002-626554/67.
DR   P-PSDB; ABP64054.
XX
PT   New polypeptide designated ORFX are present in human atherogenic cells
PT   and are useful to prevent and treat ORFX-associated disorders including
PT   cancer, allergy, wound healing or autoimmune, cardiovascular or
PT   inflammatory disease.
XX
PS   Claim 2; SEQ ID NO 847; 78pp; English.
XX
CC   The present invention relates to novel human ORFX polypeptides and their
CC   coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC   were discovered in human atherogenic cells, in particular in platelets
CC   and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC   many other tissues as well. Atherogenic cells are cells which have the
CC   potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC   nucleic acids are useful for treating or preventing a pathological
CC   condition associated with an ORFX-associated disorder, e.g. cancer,
CC   cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC   coagulation disorders or inflammatory disorders. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from the USPTO web site at
CC   seqdata.uspto.gov/sequence.html?DocID=20020082206
XX
SQ   Sequence 848 BP; 316 A; 149 C; 159 G; 223 T; 0 U; 1 Other;

Query Match   35.3%; Score 782; DB 6; Length 848;
Best Local Similarity   100.0%; Pred. No. 1.9e-130;
Matches 782; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   61 ATAACTTTTATTCTACTACTATAATGTATATGCGTGAATAGTATTAATAAATGAACTAGGGAAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 ATAACTTTTATTCTACTACTATAATGTATATGCGTGAATAGTATTAATAAATGAACTAGGGAAG 126

Qy   121 GATGTAATAAATAAGACATCTCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 GATGTAATAAATAAGACATCTCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC 186

Qy   181 TTTTCTCAAATAAATATGTGAATCAAAAGAAAAAACTTTTTTCAAGTTAATACATGGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 TTTTCTCAAATAAATATGTGAATCAAAAGAAAAAACTTTTTTCAAGTTAATACATGGT 246

Qy   241 TCAGGAAAAGAAGAAGACAAGCAAAAGAGCCAAAATCAGACTTAAGGAAAAAGAAAATAGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 TCAGGAAAAGAAGAAGACAAGCAAAAGAGCCAAAATCAGACTTAAGGAAAAAGAAAATAGA 306

Qy   301 CTTAAGTCTTCTTGTGCAGAAAACCTGAGTTCATGAAGACACCCGCTCCAGTAGATCTGGG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 CTTAAGTCTTCTTGTGCAGAAAACCTGAGTTCATGAAGACACCCGCTCCAGTAGATCTGGG 366

Qy   361 CACTTGGCCAAAGAACAAGAGTCTCCCCTGAGACGAGGCCAGTAGAATGGGGTGAATCATTT 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 CACTTGGCCAAAGAACAAGAGTCTCCCCTGAGACGAGGCCAGTAGAATGGGGTGAATCATTT 426

Qy   421 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA 486

Qy   481 TTCAGTGAAGAAATATTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487 TTCAGTGAAGAAATATTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA 546

Qy   541 CCTCAACAAATTCACCTTAAAGCAATAATGATGAAATTTATACAGACTGATGCC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547 CCTCAACAAATTCACCTTAAAGCAATAATGATGAAATTTATACAGACTGATGCC 606

Qy   601 CCAAAAGAGAGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAA 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607 CCAAAAGAGAGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAA 666

Qy   661 CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   667 CCTACCCTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC 726

Qy   721 AGTTATACACGTTTCTGAAATCTGACATCTATTAGACTTGATGGAAGAAGACCTCAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   727 AGTTATACACGTTTCTGAAATCTGACATCTATTAGACTTGATGGAAGAAGACCTCAG 786

Qy   781 AGACCAACAAATCTTAGGAGCAGATCACCGTCATTACCTGCAATGAATTCCAAGATGTA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   787 AGACCAACAAATCTTAGGAGCAGATCACCGTCATTACCTGCAATGAATTCCAAGATGTA 846

Qy   841 CA 842
          ||
Db   847 CA 848

RESULT 8
AAA52090
ID   AAA52090 standard; cDNA; 1164 BP.
XX
AC
XX
DT   04-DEC-2000 (first entry)
XX
DE   Murine RGS protein coding sequence.
XX
KW   RGS; regulators of G-protein signaling; GTPase activating protein; GAP;
KW   G-alpha protein; cell adhesion; chemotaxis; vulnerary; immunosuppressor;
KW   anti-rheumatic; anti-arthritic; anti-diabetic; anti-inflamatory;
KW   cytostatic; hepatotropic; anti-anaemic; modulator; gene therapy; ss.
XX
OS   Mus sp.
XX
FH   Key   Location/Qualifiers
FT   CDS   134..841
FT   /*tag= a
FT   /product= "RGS_protein"
XX
PN   WO20004623 6-A2.
XX
PD   10-AUG-2000.
XX
PF   04-FEB-2000; 2000WO-US002977.
XX
PR   04-FEB-1999; 99US-00244314.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Hodge MR, Yowe D;
XX
DR   WPI; 2000-532893/48.
DR   P-PSDB; AAY97154.
XX
PT   Novel regulator of G-protein signaling nucleic acids and polypeptides,
PT   useful as diagnostic and investigative tools and to treat G-protein
PT   signaling disorders.
XX
PS   Claim 1; Page 102-104; 105pp; English.
XX
CC   The RGS (regulators of G-protein signaling) protein genes, clones
CC   AAH16395 and m1975, were identified in human and murine spleen cDNA
CC   libraries, respectively. Both proteins have unique N- and C-terminal
CC   sequences. The C-terminal location of the RGS domain is consistent with
CC   RGSs known to act as GTPase activating proteins (GAPs) for G-alpha
CC   proteins. G-alpha-i linked receptors support rapid adhesion and directed
CC   migration of leukocytes and other cell types. The novel RGS proteins may
CC   be used to modulate cell adhesion and chemotaxis, e.g for aiding wound
CC   repair. The RGS proteins, related cDNAs and anti-RGS antibodies are
CC   useful for modulation, diagnosis and treatment of immune and respiratory
CC   disorders
XX

```
Qy    1062  TTTGATGTCATTCCATTATAATC 1085
            ||  |||||  |||| |||||||
Db    1055  TGTATATTCTATTCTAATAAAATTC 1078
```

RESULT 9
AAD42498
ID   AAD42498 standard; cDNA; 1164 BP.
XX
AC   AAD42498;
XX
DT   15-NOV-2002 (first entry)
XX
DE   Murine regulator of G-protein signalling (RGS) protein, m1975 cDNA.
XX
KW   Murine; screening; RGS; regulator of G-protein signalling; pancreatitis;
KW   inflammatory disorder; allergy; Grave's disease; arthritis; sinusitis;
KW   respiratory disorder; asthma; pneumonia; therapy; immune disorder;
KW   haematological disorder; haematopoiesis; platelet-associated disorder;
KW   thrombocytopaenia; migration; invasive disorder; leukaemia; anaemia;
KW   erythrocyte-associated disorder; hepatitis; m1975; gene; ss.
XX
OS   Mus sp.
XX
FH   Location/Qualifiers
FT   CDS   134..841
FT   /*tag= a
FT   /product= "Murine RGS protein"
XX
PN   US6410240-B1.
XX
PD   25-JUN-2002.
XX
PF   04-FEB-2000; 2000US-00499959.
XX
PR   04-FEB-1999; 99US-00244314.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Hodge MR, Yowe D;
XX
DR   WPI; 2002-588886/63.
DR   P-PSDB; AA825830.
XX
PT   Screening assays for identifying agent that binds to human or mouse RGS
PT   (regulators of G-protein signalling) protein or its variant, or RGS
PT   protein encoded by homologous DNA sequences, or to host cell expressing
PT   the proteins.
XX
PS   Claim 15; Col 55-58; 42pp; English.
XX
CC   The invention relates to screening assays for identifying an agent that
CC   binds to: human or mouse RGS (regulators of G-protein signalling) protein
CC   or its variant, or RGS protein encoded by homologous DNA sequences; or a
CC   host cell expressing the RGS protein or its variant, or RGS protein
CC   encoded by homologous DNA sequences. The invention is useful for
CC   identifying an agent that binds to human or mouse RGS protein or its
CC   variant, or RGS protein encoded by homologous DNA sequences, or a host
CC   cell expressing the RGS protein or its variant, or RGS protein encoded by
CC   homologous DNA sequences. The agents identified using the invention are
CC   useful for modulating the activity of RGS proteins and thus useful for
CC   treating immune and inflammatory disorders (Grave's disease, allergy,
CC   arthritis), respiratory disorders (asthma, pneumonia, sinusitis),
CC   haematological disorders (haematopoiesis, migration), platelet-associated
CC   disorders (thrombocytopaenia), invasive disorders (leukaemia),
CC   erythrocyte-associated disorders (anaemia), pancreatitis, hepatitis etc.
CC   The present sequence is murine RGS protein, m1975 cDNA
XX
SQ   Sequence 1164 BP; 406 A; 202 C; 222 G; 334 T; 0 U; 0 Other;

Query Match   26.9%; Score 596; DB 6; Length 1164;
Best Local Similarity 75.8%; Pred. No. 3.2e-97;
```

```
SQ    Sequence 1164 BP; 406 A; 202 C; 222 G; 334 T; 0 U; 0 Other;

Query Match           26.9%;  Score 596; DB 3; Length 1164;
Best Local Similarity  75.8%;  Pred. No. 3.2e-97;
Matches  791;  Conservative  0;  Mismatches 220;  Indels 33;  Gaps  3;

Qy    69  TTATTCTACTATGTATATGTATGGAATAGTATTAAATGAACTAGGGAAGGATGTAAT 128
           ||| ||  ||||||| |||| || ||||| |||||||||| | |  ||
Db    41  TTTATTCAGGATGTTTCCTATGAAATAGCATTCATCTGTGGGAGAGAAGGACTAAGG 100

Qy   129  AAATTAGACATCTCTTCAT--TTTAGAGAGAATGGAAAACAACATTGCTTTTCTTTTCT 186
           |||| ||  ||||| ||||    ||  || |  |||||| | ||| |||| | ||||
Db   101  AAATCTGACATCTGTTGGTCACTGGACAGAATATGTCACTGGTTTTTTCTTTCTCT 160

Qy   187  CAAATAAATATGTGGATCAAAAGAAAAAACTTTTTCAAGGTAATACATGGTTCAGGA 246
           ||| |||||| | |||| |||| |||||||| ||||| |||| | ||| |||||
Db   161  CAATTAAATATGTGGATCAAAAGAAAAACTTTTTTCAAACTAATGCATGGGTCAGGG 220

Qy   247  AAAGAAGAAACAAGCAAAGAGCCAAAATCAGAGCTAAGGAAAAAAAGAAATAGACTAAGT 306
           || || ||  ||| |||   |||||||||| | |||| ||||||| | | || |||||
Db   221  AAAGAAGAACAAGCATCGAGCCAAAATCAGCAGCGAAAATGCAAGAAATAGACTAAGT 280

Qy   307  CTTCTTGTGCAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTGAATCTGGGCACTTG 366
           ||| ||  |||||| |||| |||| ||| | |||| | ||| || ||| ||| |||| |
Db   281  CTTCTCCTACAGAGGCCTGACTTCCATGGAGACTCAAGCCAGTAGATCTGCCCTCTTG 340

Qy   367  GCCAAAGAAACAGAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTTGACAAA 426
           ||||||||||||| |  |||| || || |||||||| ||||| ||| |||| |||||||
Db   341  GCCAAAGAAACAAGCAAGAGTCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAA 400

Qy   427  CTGCTTCCCATAGAGATGGACTAGAGCTTTTTACCAGATTTCTTAAAACTGAATTCAGT 486
           ||||| | | || |||||| | || ||| ||||||| || ||| |||| ||||| |||
Db   401  TTGCTCTCTCATAGAGATGGATGATGCTTTTCTTACCAGATTTCTTAAAACTGAATTCAGT 460

Qy   487  GAAGAAAAATATTGAATTTTGGATAGCCTGTGAGAATTCAAGAAAAAGCAAGGGACTCAA 546
           || ||||| |||| ||||| | || | ||  | |||| ||||||||| |||| ||||| |
Db   461  GAGGAGAACAATTGAATTTTGGGTCGCCGTGAGAACTTCAAGAAAATGCAAGGAACCTCAA 520

Qy   547  CAAATTCACCTTAAAGCAAAAGCAATATATGAGAAAATTTATACAGACTGATGCCCCAAAA 606
           ||| ||||| || ||||||| | ||| | |||| ||| ||| |||| |||  || |||||
Db   521  CAAATCATCCTAAAAGCAAAAGGCAATCCTGAGAAATTCATTCAGAATGATGCCCCCAAA 580

Qy   607  GAGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAAGCAGCATCACTCAACCTACC 666
           ||||| | | |||||||||  ||  |||||  | ||  || | | |||  || ||||||||
Db   581  GAGGTTACATTGATTTTCATACTCAAAGAAGTAATTGCTAAAGCATCTGTAAGAGCCACCCACT 640

Qy   667  CTCCACAGTTTTGATGCTGCACAAAGACAGAGTGTATCAGCTCATCGGAACAAGACAGTTAT 726
           ||||||||||| |||  | | || |||||| | || | ||  | || | | |||  || |
Db   641  CTCCACAGTTTTGATACGGACACAAAGACAGAGTGTACCAGCTTACTCACGACACAGTTAT 700

Qy   727  ACACGTTTTCTGAAAATCTGACATCTATTTTGACGGAAGGAAGGGAGACCTCAGAGACCA 786
           || | ||||||||||| |||| |  | ||| ||| | |||| || |||||||||||||||
Db   701  AAACGCTTTTGAAAATCTGAGAACTCACATTTGATAGAAGAAGAGACCTCAGAGACCA 760

Qy   787  ACAAATCTTAGGAGACGATCACGCTCATTCACCTGCAATGAATTCCAAGATGTACAATCA 846
           || |||||||| || | ||||| |||||||| || || || ||||||| ||| | ||||
Db   761  ACAAACTTAGGAGGACGATCACGATCACGATCATTACTTACTTACTTACCAAGATGTAAAGTCA 820

Qy   847  GATGTTGCCATTGGTTATAAGAGAAAATGATTTTGCTCATTTTATGACAAACTTATAC 906
           ||||| ||||||||||||| | || ||| |||||||| || ||| |||| | |||
Db   821  GATGTTGCCATTGGTTATGAGTAAAAGTCATTTGTCTTTCTTTTTGATAGTGTATGTGTAT 880

Qy   907  ATC-------------------TGCTTCTAACATATCGCATGTTTATGTTAAGA 941
           |||                  ||  || ||||  | | ||| |  |||||| |
Db   881  ATTCTAAAATATATACTAATACTAATGTGTACTTCTAAAATATAGCTTGTGTATAAGAGA 940

Qy   942  TTTGGTCCCATCCTTAAACTGCAAATAATGTCATGTGAAATTATTTTAAAAATGTAAAAAC 1001
           ||||  | || | |||       |||| ||| || || || ||||| ||||||||||| |
Db   941  GATGATTTCATTTT-------TAAAATACCACCATGCAAATAACATATTAAAATGTAAGACTT 994

Qy  1002  AAAACTTTCTGCTAACAAAATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTA 1061
           |||| ||| | ||  |  | | | |     ||| || |||| |||  ||||||| |||
Db   995  TTTATATTATACTAAATAATTCATCATCTATCTTCCGAAATATTTTATGAAAATCTATC 1054
```

```
RESULT 10
AAS18334
ID   AAS18334 standard; cDNA; 241 BP.
XX
AC   AAS18334;
XX
DT   12-MAR-2002 (first entry)
XX
DE   cDNA encoding partial human platelet RGS domain.
XX
KW   Human; regulator of G protein signalling; RGS18; arterial thrombosis;
KW   platelet activation dysfunction; myocardial infarction; stroke;
KW   coronary artery disease; cerebrovascular disease; unstable angina;
KW   deep vein thrombosis; systemic thromboembolism; anti-coagulant;
KW   invasive cardiac procedure; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..240
FT                   /tag= a
FT                   /partial
FT                   /product= "Partial RGS domain"
FT                   /note= "This sequence encodes for residues 2-81 of
FT                   AAU10747 and lacks both start and stop codons"
XX
PN   WO200183514-A2.
XX
PD   08-NOV-2001.
XX
PF   26-APR-2001; 2001WO-US013540.
XX
PR   28-APR-2000; 2000US-0200786P.
PR   02-AUG-2000; 2000GB-00018833.
XX
PA   (AVET ) AVENTIS PHARM PROD INC.
XX
PI   Murray DL, Gagnon AW;
XX
DR   WPI; 2002-055453/07.
DR   P-PSDB; AAU10747.
XX
PT   Isolated regulator of G protein signaling polypeptide, useful for
PT   prevention/treatment of platelet activation dysfunction, such as arterial
PT   thrombosis, myocardial infarction, coronary artery disease and stroke.
XX
PS   Claim 1; Page 115; 127pp; English.
XX
CC   The present invention relates to the isolation of polynucleotide
CC   sequences, that encode a novel regulator of G protein signalling (RGS)
CC   polypeptide, RGS18, from human platelets. The invention also provides
CC   nucleotide primers and probes specific for an RGS18 nucleic acid. The
CC   sequences of the invention are useful for the manufacture of a medicament
CC   for the prevention or treatment of a platelet activation dysfunction,
CC   such as arterial thrombosis, myocardial infarction, coronary artery
CC   disease, stroke, cerebrovascular disease, unstable angina, deep vein
CC   thrombosis, systemic thromboembolism, and in invasive cardiac procedures
CC   for anti-coagulant purposes. The sequences of the invention can also be
CC   used for the detection of RGS18 nucleic acids and RGS18 polypeptides. The
CC   present cDNA sequence encodes for a partial human platelet RGS domain
XX
SQ   Sequence 241 BP; 92 A; 52 C; 44 G; 53 T; 0 U; 0 Other;
```

```
Query Match             10.2%;  Score 227.2;  DB 6;  Length 241;
Best Local Similarity   96.7%;  Pred. No. 1.9e-31;
Matches 232;   Conservative    0;   Mismatches   8;   Indels   0;   Gaps   0;

Qy       487 GAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGACCTCAA 546
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GAGGAAAATCTGAGTTCTGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGACCTCAA 60
```

Matches 791;   Conservative   0;   Mismatches 220;   Indels   33;   Gaps   3;

```
Qy    69 TTATTCTACTATGTATATGTGGAATAGTAGTAATAAATAGAACTAGGGAAGGATGTAAT 128
         |||||  ||  || |||| ||  ||||  ||||  ||||||| |||||||||||||||
Db    41 TTTATTCAGGATGTTTTCCTATGAAATAGCATTCATCTGGGAGAGAAGGACTAAGG 100

Qy   129 AAATTAGACATCTCCTTCAT--TTTAGAGGATGGAAACAACATTGCTTTCTTTTCT 186
         |||| ||||||||| |||   ||| |||| |||| ||||| || ||||| |||||
Db   101 AAATCTGACATCTGTTGGTCACTGGTTTTCTCTGTTCTCTCT 160

Qy   187 CAAATAAATATGTGAATCAAAAGAAAAACTTTTTTCAAGTAATACATGGTTCAGGA 246
         |||| ||||||||||||||||| |||||||| ||||||||| |||| |||||||||
Db   161 CAATTAAATATGTGAATCAAAAGAGAAAAACTTTTTTCAAACTAATGCATGGTCAGGG 220

Qy   247 AAAGAAGAAACAAGCAACCAAATCAGAGCTAAGGAGAAAAAGAAATGAAATGACTAAGT 306
         |||| |||||||||| |||| ||||| ||||||||||| |||||||||||||||||||
Db   221 AAAGAAGAAACAAGCAACCATCGAGGCCAAAATCAGAGGCCAAAGAAGAAATGAGACTAAGT 280

Qy   307 CTTCTTGTGCAGAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGAGTCTGGGCACTTG 366
         ||||  ||||||||||||||| |||||||||||| ||||||| ||||| |||||||||
Db   281 CTTCCTACAGAGGGCCTTGACTTCCATGGAGAGACTAGAAGACTGATCTGCCCTCTTG 340

Qy   367 GCCAAGAAACAAGAGTCTCCCCTGAAGAGGGCAGTGAAATGGGGTGAATCATTTGACAAA 426
         |||||||||| ||||| | | |||||||||||| |||||| ||  |||||||||||||||
Db   341 GCCAAGAAACAAGAGTCTCTCGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAA 400

Qy   427 CTGCTTTCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 486
         ||||| | ||| |||||||||| ||||| || |||||||||||| |||||||||||||
Db   401 TTGCTCTCCATAGAGATGGGAGTGCTTTTCACCAGATTTCTTAAAACTGAATTCAGT 460

Qy   487 GAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGACCTCAA 546
         ||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db   461 GAGGAGAACAATTGAATTTGGGTCGCCGTGTGAAGACTTCAAGAAATGCAAGGAACCTCAA 520

Qy   547 CAAATTCACCTTAAAGACAAAAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAA 606
         ||||| |||| |||||||||| |||||| |||||| |||||||||||||||||||||||
Db   521 CAAATCATCCTAAAAGCAAGGCAATCATGAGAAATTCATTCAGAATGATGCCCCCAAA 580

Qy   607 GAGGTTAACCTTGATTTTCACACAGAAGAGTCATTACAAACAGCATCACTCAACCTACC 666
         ||||||| ||||||||||||| ||||||||| |||||| ||||||||||||||||||||
Db   581 GAGGTTAACATTGATTTTCATCTCAAGAGAAGTAATTGCTAACGAGCATCGCCCAGCCCACT 640

Qy   667 CTCCACAGTTTGTGCTGCACAAAGCAGAGTCTCATGGAACAAGACAGTTAT 726
         |||||||||| || || |||||||||||||| ||| ||||||||||||||||
Db   641 CTCCACAGTTTTGATACGGCACCACAAAGCAGAGTGTACCAGCTCCATCGACAGTTAT 700

Qy   727 ACACGTTTTCTGAAATCTGACATCTGATTAGACTTGAGGATGGAGGAAGACCTCAGAGACCA 786
         ||||||||| ||||||||||| || | |||| | || ||| |||||||||||||||||||
Db   701 AAACGCTTTTTGAAATCTGAGACGTTACTTACATTTGATAGAGGAGACCTCAGAGACCA 760

Qy   787 ACAAATCTTAGGAGGAGGATCACGTCATTTACCTGCAATGAATTCCAAGATGTACAATCA 846
         |||| | ||||||||||||||||||| |||||||||| || | || ||||||||||||||
Db   761 ACAAACCTTAGGAGGACGACGATCACGATCATCATTACTTACTTACACAATGTAAAGTCA 820

Qy   847 GATGTTGCCATTTGGTATAAAGAAATTGATTTGCTCATTTTTGATAGTGTATGTGTAT 906
         |||||||||||||||||||||| |||||||||| | |||||  ||||| ||||||||||
Db   821 GATGTTGCCATTTGGTTATGAGTAAAAGTCATTTCGTCTTTCTTTGATAGTGTATGTGTAT 880

Qy   907 ATC-----------------TGCTTCTAACATATCGCATGTTTATGTTAAGA 941
         |||                 |||||||||||||| |||| |||||||||||||
Db   881 ATCTAAAATATATACTAATGTGTACTTCTAAAATATAGCTTGTGTATAAGAGA 940

Qy   942 TTTGGTCCCATCCTTAAACTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAAC 1001
         |||||||| |||| ||||  |||| |||||||||||||||||||||||||||||||||||
Db   941 GATGATTTCATTTT-----TAAAATACCACCATGCAAATACAATTACATATTAAAATGTAAGACTT 994

Qy  1002 AAAACTTTCTGCTAACAAAATCACTACAGTATCGCCAGTATATCGTAAAACCTTCTA 1061
         |||||||| | ||| || |||||||||| ||| ||| ||||| | ||||||||||||
Db   995 TTTATTATTACTAAAATAATTCATCATCATCTTCCGAAATATTTTATGAAAATCTATC 1054

Qy  1062 TTTGATGTCATTCCATTATAAATC 1085
         |||||| |||| |||||| |||||
Db  1055 TGATATTCTATTCTATTCAATAAAATTC 1078
```

```
Qy   547  CAAATTCACCTTAAAGCAAAAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAA  606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CAAATTCACCTTAAAGCAAAAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAA  120

Qy   607  GAGGTTAACCTTGATTTCACACAAAGAGAAGTCATTACAACAGCATCACTCAACCTACC  666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GAGGTTAACCTTGATTTCACACAAAGAGAAGTCATTACAACAGCATCACTCAACCTACC  180

Qy   667  CTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGACAAGACAGTTAT  726
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  CTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAAAACGCAGCTAT  240
```

RESULT 11
ACC43243
ID   ACC43243 standard; DNA; 408 BP.
XX
AC   ACC43243;
XX
DT   11-AUG-2003 (first entry)
XX
DE   Nucleotide sequence of regulator polypeptide of G-protein signalling.
XX
KW   Human; regulator polypeptide of G-protein signalling; RGS; chromosome 1;
KW   asthma; diabetes; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   4..408
FT   /*tag= a
FT   /product= "regulator polypeptide of G-protein signalling"
XX
PN   WO200301634S-A1.
XX
PD   27-FEB-2003.
XX
PF   16-AUG-2002; 2002WO-EP009173.
XX
PR   16-AUG-2001; 2001US-0312345P.
XX
PA   (FARB) BAYER AG.
XX
PI   Koehler RH;
XX
DR   WPI; 2003-278547/27.
DR   P-PSDB; ABP98021.
XX
PT   New polynucleotide encoding a regulator polypeptide of G-protein
PT   signaling (RGS) and RGS protein, useful in gene therapy to modulate the
PT   activity of RGS in a disease, particularly in asthma or diabetes.
XX
PS   Claim 1; Fig 4; 99pp; English.
XX
CC   The present sequence encodes an alternatively spliced version of human
CC   regulator polypeptide of G-protein signaling (RGS). The RGS gene is
CC   located on chromosome 1. The RGS polynucleotide and polypeptide are
CC   useful for identifying test compounds, that may act as agonists or
CC   antagonists at the receptor site and which can be regulated to provide
CC   therapeutic effects. The polynucleotide, polypeptide or vectors
CC   comprising the polynucleotide, and the reagents are useful for modulating
CC   the activity of RGS in a disease, particularly in asthma or diabetes
XX
SQ   Sequence 408 BP; 139 A; 77 C; 77 G; 115 T; 0 U; 0 Other;

Query Match   7.1%;   Score 158.4;   DB 7;   Length 408;
Best Local Similarity   63.8%;   Pred. No. 3.8e-19;
Matches   240; Conservative   0; Mismatches   136; Indels   0; Gaps   0;

```
Qy   389  CTGAAGAGGCAGTGAAATGGGGTGAATCATTTGACAAACTGCTTTCCCATAGAGATGGAC  448
          |||||||    ||| |||||||  |  |||||||||| ||||||||||||||||||||||
Db    17  CTGCGGAAACAATGACATGGTCTGAAAATATCGGACACGCTTTTTAGCCAACCAAGCTGGTC  76
```

```
Qy   449  TAGAGGCTTTCACCAGATTCTTCTTAAAACTGAATTCAGTGAAGAAAAATATTGAATTTGGA  508
          |||| || |||||||||||||||||  ||||||||||||||||||||||||||||||||
Db    77  TAGATGCTTTTCGAATATTTCTAAAATCAGAGTTCTAAAATCAGAGTTAGTGAAGAAAATGTTGAGTTCTGGC  136

Qy   509  TAGCCTGTGAAGATTCAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAAAGCAAAAG  568
          ||||||||||||  |||| |||||| |||  ||||||||||||||||||||||||||||
Db   137  TTGCCTGTGGAGACTTTAAGAAAACGAAAAATCGAGACAAAATTGCTTCCAAAGCCAGA  196

Qy   569  CAATATATGAGAAATTATACAGACTGATGCCCCAAAAGAGGTTAACCTTGATTTTCACA  628
          |||||||| ||||||| ||| |||||||| || |||||||| || ||  ||||||||
Db   197  TGATTATTCTGAATTCATTGAGCTGATGCCACCTCTGTTGTGCACCTTAACATTGACTTCGGTA  256

Qy   629  CAAAAGAGTCATTACAACAGCATCACTCAACCTACCTCCACAGTTTGATGCTGCAC  688
          ||| |||||||  |   | ||| || |||||| ||  |||||||||||||||
Db   257  CCAGAGACCTCATCTCAAGAATATTGCTGAACAAGAGAGTTATACACGTTTATACAGCTTTGAAATCTGACA  316

Qy   689  AAAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTATACACGTTTTCTGAAATCTGACA  748
          |  ||  ||| |||  ||||| || |||  |||| | |||  | |||| |||||| ||||
Db   317  AGAAATTAATCTATTGTCTCATGGCCAAGGATTCTTTCCCTCCGATTTCTGAAGTCAGAGA  376

Qy   749  TCTATTTAGACTTGAT  764
          ||| ||||  ||
Db   377  TTTATAAAAACTGGT  392
```

RESULT 12
AAH76414
ID   AAH76414 standard; cDNA; 1795 BP.
XX
AC   AAH76414;
XX
DT   04-OCT-2001 (first entry)
XX
DE   RGS protein 17 cDNA.
XX
KW   RGS protein 17; cytostatic; virucidal; immunomodulatory;
KW   antiinflammatory; haemostatic; cancer; haemopathy; infection;
KW   human immunodeficiency virus; HIV; immunological disease;
KW   inflammatory disease; ss.
XX
OS   Unidentified.
XX
PN   WO200155414-A1.
XX
PD   02-AUG-2001.
XX
PF   15-JAN-2001; 2001WO-CN000032.
XX
PR   26-JAN-2000; 2000CN-00111515.
XX
PA   (BIOD-) BIODOOR GENE TECHNOLOGY LTD SHANGHAI.
XX
PI   Mao Y, Xie Y;
XX
DR   WPI; 2001-476216/51.
DR   P-PSDB; AAG66509.
XX
PT   RGS protein 17 and encoded polynucleotide, applicable in diagnosis and
PT   treatment of malignant tumor, hemopathy, HIV infection, immunological
PT   diseases and various inflammations.
XX
PS   Claim 6; Page 29-30; 35pp; Chinese.
XX
CC   The invention relates to an isolated polypeptide of RGS protein 17
CC   comprising the 152 amino acid sequence defined in the specification, or
CC   its fragment, analogue or derivative. The polypeptide and the
CC   polynucleotide encoding it are useful in the diagnosis and treatment of
CC   malignant tumours, haemopathy, HIV infection, immunological diseases and
CC   various inflammatory diseases. The present sequence encodes the
CC   polypeptide of the invention
XX
SQ   Sequence 1795 BP; 632 A; 295 C; 281 G; 587 T; 0 U; 0 Other;

CC   ABT1112, where a differential expression of the genes indicates breast
CC   cancer. The methods are useful in diagnosing, treating, detecting the
CC   progression, and in monitoring treatment of breast cancer in patients.
CC   The methods are also useful as a screening tool for agents that modulate
CC   the onset or progression of breast cancer. The breast cancer genes may be
CC   used as diagnostic markers for the prediction or identification of the
CC   malignant state of breast tissue, for confirming the type and progression
CC   of cancer, and for drug screening and assays. The present sequence is a
CC   coding sequence of the invention. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub.published_pct_sequences
XX
SQ   Sequence 1345 BP; 443 A; 254 C; 304 G; 344 T; 0 U; 0 Other;

Query Match   7.0%; Score 155.6; DB 6; Length 1345;
Best Local Similarity 53.2%; Pred. No. 1.4e-18;
Matches   239; Conservative   0; Mismatches 139; Indels   0; Gaps   0;

Qy   385 TCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTTGACAAACTGCTTTCCATAGAGAT 444
            |||||| ||| ||||||| |||||| |||| ||||||||| || ||||| |||||
Db   249 TCTCCCTGAGGAAGCACAGCTGTGGTCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy   445 GGACTAGAGGCTTTACCAGATTTCTTAAAACTGAATTCAGTGAAGAAAAATATTGAATT 504
            || |||||||||| |||||| ||||| ||||| ||| || || |||| ||||||||||||
Db   309 GGTCTTGCTGCATTCAGGGCTTTTTTAAAGTCGGAATTCGTGTGAAGAAAAATATTGAATTC 368

Qy   505 TGGATAGCCTGTGAAGATTCAAGAAAGAAGAGGACCTCAACAAATTCACCTTAAAGCA 564
            ||||| |||||||||||| |||| ||||||| |||| ||| ||||| | ||||| ||
Db   369 TGGCTGCCTGTGAAGACTTCAAAAAATTATACAGACTGATGCCCCAAAAGAGGTTAACCTGATTT 428

Qy   565 AAAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAAGAGGTTAACCTTGATTTT 624
            |||| |||||||||||| | ||||||| |||||||||||||||| ||||||||||||||
Db   429 AGGAAAATATATACTGACTTCATAGAAAAGGAAGCTCCAAAAGAATAAACATAGATTTT 488

Qy   625 CACACAAAGAGAGTCATTACAAACAGCATCTAACCTACCCTCCACAGTTTTGATGCT 684
            |||||||| ||| ||| |||| ||||| || |||| ||| ||||||| |||||||||
Db   489 CAAACCAAAACTCTGATTGCCCAGAATATACAAGAGCTACAAGTGGCTGCTTTACAACT 548

Qy   685 GCACAAAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTATACACGTTTTCTGAAAATCT 744
            || |||||| |||||||||||| ||||||| || ||||||||| ||||| ||||||||
Db   549 GCCCAGAAAGGGTATACAGCTTGATGGAGAACAACTCTTATCCTCGTTTCTTCTTGGAGTCA 608

Qy   745 GACATCATTTTAGACTTG 762
            || ||||| ||||||||
Db   609 GAATTCTTACCAGGACTTG 626

RESULT 14
ABK83834
ID   ABK83834 standard; cDNA; 1345 BP.
XX
AC   ABK83834;
XX
DT   14-AUG-2002   (first entry)
XX
DE   Human cDNA differentially expressed in granulocytic cells #405.
XX
KW   Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW   viral infection; parasitic infection; protozoal infection;
KW   fungal infection; sterile inflammatory disease; psoriasis;
KW   rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW   cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW   adult respiratory distress syndrome; inflammatory bowel disease;
KW   Crohn's disease; ulcerative colitis; periodontal disease;
KW   granulocyte activation; chronic inflammation; allergy.
XX
OS   Homo sapiens.
XX
PN   WO200228999-A2.
XX
PD   11-APR-2002.
XX

Query Match   7.1%; Score 158.4; DB 4; Length 1795;
Best Local Similarity 63.8%; Pred. No. 4.4e-19;
Matches   240; Conservative   0; Mismatches 136; Indels   0; Gaps   0;

Qy   389 CTGAAGAGGCAGTGAAATGGGTGAATCATTTGACAAACTGCTTTCCATAGAGATGGAC 448
            ||||| | ||||||| |||||| |||| ||||||||| || ||||| |||||||
Db   210 CTGCGGAAACAATGACATGGTCTGAAAATATCGACACGCTTTTAGCCAACCAAGCTGGTC 269

Qy   449 TAGAGGCTTTACCAGATTTCTTAAAACTGAATTCAGTGAAGAAAAATATTGAATTTTGGA 508
            |||||||||| |||||| ||||| ||||| ||| || || |||| ||||||||||||
Db   270 TAGATGCTTTTCGAATATTTCTAAAATCAGAGTTGAAGAAAAATGTTGAGTTCTGGC 329

Qy   509 TAGCCTGTGAAGATTCAAGAAAGAAGAGGGACCTCAACAAATTCACCTTAAAGCAAAAG 568
            ||||||||||||| |||| ||||||||| ||| ||| ||||| | ||||| ||
Db   330 TTGCCTGTGAAGACTTTAAGAAAACAGAAAATGCGACAGCAAATTGCTTCCAAAGCCAAGA 389

Qy   569 CAATATATGAGAAATTTATACGACTGATGCCCCAAAAGAGGTTAACCTTGATTTTCACA 628
            ||||||||||||| | |||||| ||||||||||||||||| |||||||||||||||
Db   390 TGATTTATTCTGAATTCATTGAAGCTGATGCACCTACCCTCCACAGTTTTGATGTGCAC 449

Qy   629 CAAAAGAAGTCATTACAAACAGCATCAACTCAACCTACCCTCCACAGTTTTGATGTGCAC 688
            |||||| ||| ||| |||| |||| | |||| ||| ||||||| |||||||||||||
Db   450 CCAGAGAACCTCCATCCTCAAAGAGAATATTGCTGAACCAACACTCAAATGCTTTGATGAGGCTC 509

Qy   689 AAAGCAGAGAGTGTATCAGCTCATGCAACAGACAGTTATACACGTTTTTCTGAAAATCGACA 748
            |||| | |||||||||||||| |||||| |||| ||||||||| ||||| |||||||
Db   510 AGAAATTAATCTATTGTCTCATGGCCCAAGGATTCTTTCCCCTGGATTTTCTCGATTTCTGGAGTGA 569

Qy   749 TCTATTTAGACTTGAT 764
            ||||||||||||
Db   570 TTTATAAAAACTGGT 585

RESULT 13
ABT10881
ID   ABT10881 standard; cDNA; 1345 BP.
XX
AC   ABT10881;
XX
DT   04-DEC-2002   (first entry)
XX
DE   Human breast cancer associated coding sequence SEQ ID NO: 1015.
XX
KW   Human; breast specific gene; breast cancer; differential expression;
KW   cytostatic; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200259271-A2.
XX
PD   01-AUG-2002.
XX
PF   25-JAN-2002; 2002WO-US002176.
XX
PR   25-JAN-2001; 2001US-0263757P.
PR   25-APR-2001; 2001US-0286090P.
PR   23-MAY-2001; 2001US-0292517P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Orr MS, Nation M, Diggans JC, Zeng W;
XX
DR   WPI; 2002-674803/72.
XX
PT   Diagnosing breast cancer in a patient comprises detecting the level of
PT   gene expression in cell or tissue samples, where a differential gene
PT   expression is indicative of breast cancer.
XX
PS   Claim 1; SEQ ID NO 1015; 260pp + Sequence Listing; English.
XX
CC   The present invention relates to methods of diagnosing breast cancer in a
CC   patient, which comprise detecting the level of expression in a tissue
CC   sample of two or more genes selected from those shown in ABT09867-

PF   03-OCT-2001; 2001WO-US030821.
XX
PR   03-OCT-2000; 2000US-0237189P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR   WPI; 2002-435328/46.
XX
PT   Detecting granulocyte activation by detecting differential expression of
PT   genes associated with granulocyte activation, which serves as diagnostic
PT   markers that is useful for monitoring disease states and drug toxicity.
XX
PS   Claim 1; SEQ ID NO 405; 114pp; English.
XX
CC   The invention relates to detecting the level of expression (M1) granulocyte (GC) activation
CC   (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC   DNA chip analysis as given in the specification, and comparing the
CC   expression level to an expression level in an unactivated GC, where
CC   differential expression of Gs is indicative of GCA. Also included are
CC   modulating (M2) GA by contacting GC with an agent that alters the
CC   expression of at least one gene in Gs; (2) screening (M3) for an agent
CC   capable of modulating GCA or an inflammation (especially chronic) in a
CC   tissue, an allergic response in a subject, exposure of a subject to a
CC   pathogen or sterile inflammatory disease using the gene expression
CC   profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC   tissue, an allergic response in a subject, exposure of a subject to a
CC   pathogen or sterile inflammatory disease, by detecting the level of
CC   expression in a sample of the tissue of gene(s) from Gs, where the level
CC   of expression of the gene is indicative of inflammation; (4) treating
CC   (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC   response in a subject, exposure of a subject to a pathogen or sterile
CC   inflammatory disease, by contacting a tissue having inflammation with an
CC   agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC   is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC   for screening an agent capable of modulating GCA preferably in an
CC   inflammation in a tissue; M4 is useful for detecting an inflammation
CC   (especially chronic) in a tissue, an allergic response in a subject,
CC   exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC   psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC   cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC   respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC   disease, ulcerative colitis, periodontal disease; also bacterial
CC   infection, viral infection, parasitic infection, protozoal infection,
CC   fungal infection and M5 is useful for treating one of the above
CC   conditions. The present sequence represents a gene differentially
CC   expressed in granulocytes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1345 BP; 443 A; 254 C; 304 G; 344 T; 0 U; 0 Other;

Query Match   7.0%;   Score 155.6;   DB 6;   Length 1345;
Best Local Similarity 63.2%;   Pred. No. 1.4e-18;
Matches 239;   Conservative   0;   Mismatches 139;   Indels   0;   Gaps   0;

Qy   385 TCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTGACAAACTGCTTCCCATAGAGAT 444
            || |||||||||||| |||||| |||| || ||||||||| ||| |||||||||||
Db   249 TCTCCTGGAGAAGCACAGCTGTGGTCAGAAGCATTTGACGAGCTGCTGACCAGCAAATAT 308

Qy   445 GGACTAGAGGCTTTACCAGATTTCTTAAAACTGAATTCAGTGAAGAAAATATTGAATTT 504
            ||||||||||||||||||| ||| ||||||||||||||||| |||||| || |||||||
Db   309 GGTCTTGCTGCATTCAGGGCGTTTTTTAAAGTCGGAATTCTGTGAAGAAAATATTGAATTC 368

Qy   505 TGGATAGGCTGTGAAGATTTCAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAAAGCA 564
            ||| ||||||||||||||||||| ||||| |||| |||||||||||| |||||||||||
Db   369 TGGCTGGCTGTGAAGAACTTCAAAAAAAAACCAAATCACCCCAAAAGCTGTCCTCAAAAGCA 428

Qy   565 AAAGCAATATATGAGAAATTATACAGACTGATGCCCCAAAAGAGAGTTAACCTTGATTTT 624
            |||| |||||||| |||||||||||||| |||||||| ||| |||||||||||||||
Db   429 AGGAAAATATATACTGACTTCATGAGAAAAAGAAGCTCCAAAAGAGATAAAACTAGATTTT 488

Qy   625 CACACAAAAGAGTCATTACAAAACAGCATCACTCAACCTACCTCACCTCCACAGTTTGATGCT 684
            ||||||||| ||| ||| ||||||||||||||||||| ||||| ||| ||||| |||||||
Db   489 CAAACCAAAACTCTGATTGCCCAGAATACAAGAAGCTACAAGTGGCTGCTTACAACT 548

Qy   685 GCACAAAGCAGAGTGTATCAGCTCATGGAACAAGAACAGTTATACAGTTTTCGAAATCT 744
            || || ||||||||||||| ||||||||| || || ||||||||||||| ||||||||
Db   549 GCCCAGAAAAGGGTACACAGTTGATGGAGACAAACTCTTATCCTCGTTTCTTGGAGTCA 608

Qy   745 GACATCTATTTAGACTTG 762
            || |||| |||||
Db   609 GAATTCTACCAGGACTTG 626

RESULT 15
ACC46750
ID   ACC46750   standard; cDNA; 1345 BP.
XX
AC   ACC46750;
XX
DT   05-JUN-2003   (first entry)
XX
DE   Human COPD related protein encoding cDNA SEQ ID NO:1.
XX
KW   Human; chronic obstructive pulmonary disease; COPD; chronic lung disease;
KW   gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO20027127-A2.
XX
PD   05-DEC-2002.
XX
PF   28-MAY-2002; 2002WO-EP005835.
XX
PR   31-MAY-2001; 2001GB-0001 3266.
XX
PA   (FARB) BAYER AG.
XX
PI   Oellers N, Gehrmann M, Kallabis H, Hall R, Schulze T, Kroegel C;
XX
DR   WPI; 2003-140492/13.
DR   P-PSDB; ABP96779.
XX
PT   Predicting, diagnosing or prognosing chronic lung disease, by detecting a
PT   chronic obstructive pulmonary disease (COPD) gene in a biological sample.
XX
PS   Claim 8; Page 73-74; 214pp; English.
XX
CC   The present invention describes a method for predicting, diagnosing or
CC   prognosing chronic lung disease by detecting a chronic obstructive
CC   pulmonary disease (COPD) gene related polynucleotide (see ACC46750 to
CC   ACC46777, which encode the COPD related proteins in ABP96779 to
CC   ABP96806). The method is useful for predicting, diagnosing or prognosing
CC   chronic lung disease in a biological sample. The COPD genes and proteins
CC   encoded by them from the present invention (I) can be used for treating
CC   or preventing chronic lung disease in a mammal. (I) can be used in an
CC   animal model for determining the efficacy, toxicity, or side effects of
CC   treatment with (I), and determining the mechanism of action of (I).
CC   ACC46778 to ACC46903 represent COPD related PCR primers and probes used
CC   in an example from the present invention
XX
SQ   Sequence 1345 BP; 443 A; 254 C; 304 G; 344 T; 0 U; 0 Other;

Query Match   7.0%;   Score 155.6;   DB 7;   Length 1345;
Best Local Similarity 63.2%;   Pred. No. 1.4e-18;
Matches 239;   Conservative   0;   Mismatches 139;   Indels   0;   Gaps   0;

Qy   385 TCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTGACAAACTGCTTCCCATAGAGAT 444
            || |||||||||||| |||||| |||| || ||||||||| ||| |||||||||||
Db   249 TCTCCTGGAGAAGCACAGCTGTGGTCAGAAGCATTTGACGAGCTGCTGACCAGCAAATAT 308

Qy   445 GGACTAGAGGCTTTACCAGATTTCTTAAAACTGAATTCAGTGAAGAAAATATTGAATTT 504

```
Db        |||  ||  |  |||  |  |||  |||  ||   |  |||||  ||||  ||||||||||||||||||
      309 GGTCTTGCTGCATTCAGGGCTTTTTAAAGTCGGAATTCTGTGAAGAAAATATTGAATTC 368

Qy        ||  |||||||||  |||||||  |  ||||||||||||||||  |  ||  ||||||
      505 TCGGATAGCCCTGTGAAGATTTCAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAAAGCA 564
Db    369 TGGCTGGCCTGTGAAGACTTCAAAAAAACCAAATCACCCCAAAAAGCTGTCCTCAAAAGCA 428

Qy    565 AAAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAAGAGGTTAACCTTGATTTT 624
          |  ||||||  |||  |  |||  |  ||||  |||||  |||||||||  |  |||||
Db    429 AGGAAAATATATACTGACTTCATAGAAAAGGAAGCTCAAAAAGAGATAAACATAGATTTT 488

Qy    625 CACACAAAAGAGAGTCATTACAAACAGCATCACTCAACCTCACCCTCCACAGTTTTGATGCT 684
          ||  |||||  |||||  |  |  |||||  ||  |||||
Db    489 CAAACCAAAACTCTGATTGCCCAGAATATACAAGAAGCTACAAGTGGCTGCTTTACAACT 548

Qy    685 GCACAAAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTATACACGTTTTCTGAAATCT 744
          ||  ||||||  |||  |  |||||  |  ||||||||  ||  ||  ||
Db    549 GCCCAGAAAAGGGTATACAGCTTGATGGAGAACAACTCTTATCCCTCGTTTCTTGGAGTCA 608

Qy    745 GACATCTATTTAGACTTG 762
          ||  ||||  ||||||
Db    609 GAATTCTACCAGGACTTG 626
```

Search completed: August 20, 2004, 06:15:33
Job time : 928.824 secs

Copyright GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2004, 05:14:44 ; Search time 175.734 Seconds
                                           (without alignments)
                                           7001.078 Million cell updates/sec

Title: US-09-894-749-1
Perfect score: 2217
Sequence: 1 gaattcggcttccatcccaa........aaaaaaaaaggggcggccgc 2217

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued Patents NA:*
               1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
               2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
               3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
               4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
               5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
               6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Query Length	DB	ID	Description
1	2217	100.0	2217	3	US-09-244-314-1	Sequence 1, Appli
2	2217	100.0	2217	3	US-09-499-959-1	Sequence 1, Appli
3	596	26.9	1164	3	US-09-244-314-3	Sequence 3, Appli
4	596	26.9	1164	4	US-09-498-959-3	Sequence 3, Appli
5	388	17.5	736	4	US-09-833-381-1010	Sequence 1010, Ap
6	155.6	7.0	1345	4	US-09-016-434-1278	Sequence 1278, Ap
7	155.6	7.0	1345	4	US-09-023-655-1243	Sequence 1243, Ap
8	153.2	6.9	744	2	US-08-748-483-2	Sequence 2, Appli
9	153.2	6.9	2190	4	US-09-023-655-795	Sequence 795, App
10	146.6	6.6	3208	4	US-09-566-921-111	Sequence 111, App
11	143.8	6.5	800	4	US-09-016-434-1091	Sequence 1091, App
12	141.2	6.4	1398	4	US-09-016-434-1290	Sequence 1290, App
13	141.2	6.4	1398	4	US-09-702-705-322	Sequence 322, App
14	141.2	6.4	1398	4	US-09-736-457-322	Sequence 322, App
15	141.2	6.4	1398	4	US-09-614-124B-322	Sequence 322, App
16	141.2	6.4	1398	4	US-09-671-325-322	Sequence 322, App
17	141.2	6.4	1398	4	US-09-589-184-322	Sequence 322, App
18	127	5.7	2638	4	US-09-016-434-1090	Sequence 1090, Ap
19	115.8	5.2	606	2	US-08-463-081B-27	Sequence 27, Appli
20	115.8	5.2	606	2	US-08-461-379A-27	Sequence 27, Appli
21	115.8	5.2	606	2	US-08-462-390B-27	Sequence 27, Appli
22	115.8	5.2	606	3	US-08-463-074B-27	Sequence 27, Appli
23	115.8	5.2	606	3	US-08-465-585C-27	Sequence 27, Appli
24	115.8	5.2	606	3	US-08-652-446-27	Sequence 27, Appli
25	115.8	5.2	746	3	US-08-870-815-3	Sequence 3, Appli
26	115.8	5.2	746	3	US-08-949-004-3	Sequence 3, Appli
27	115.8	5.2	2383	1	US-08-274-318-1	Sequence 1, Appli
28	115.8	5.2	2383	2	US-08-754-108-1	Sequence 1, Appli
29	115.8	5.2	2406	2	US-08-463-081B-1	Sequence 1, Appli
30	115.8	5.2	2406	2	US-08-461-379A-1	Sequence 1, Appli
31	115.8	5.2	2406	3	US-08-462-390B-1	Sequence 1, Appli
32	115.8	5.2	2406	3	US-08-463-074B-1	Sequence 1, Appli
33	115.8	5.2	2406	3	US-08-465-585C-1	Sequence 1, Appli
34	115.8	5.2	2406	2	US-08-652-446-1	Sequence 1, Appli
35	107.2	4.8	981	2	US-08-829-110-2	Sequence 2, Appli
36	107.2	4.8	981	4	US-09-016-434-240	Sequence 240, App
37	95.4	4.3	1225	2	US-08-829-110-4	Sequence 4, Appli
38	95.4	4.3	2272	2	US-08-726-228-1	Sequence 1, Appli
39	95.4	4.3	2272	3	US-08-870-815-1	Sequence 1, Appli
40	95.4	4.3	2272	3	US-08-949-004-1	Sequence 1, Appli
41	91.2	4.1	498	0	US-09-702-705-42	Sequence 42, Appl
42	91.2	4.1	498	0	US-09-736-457-42	Sequence 42, Appl
43	91.2	4.1	498	0	US-09-614-124B-42	Sequence 42, Appl
44	91.2	4.1	498	0	US-09-671-325-42	Sequence 42, Appl
45	91.2	4.1	498	0	US-09-589-184-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-09-244-314-1
; Sequence 1, Application US/09244314
; Patent No. 6274362
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; APPLICANT: Yowe, David
; TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
; FILE REFERENCE: 5800-19, 035800/174680
; CURRENT APPLICATION NUMBER: US/09/244,314
; CURRENT FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (160)..(867)
US-09-244-314-1

Query Match   100.0%; Score 2217; DB 3; Length 2217;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 2217 Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 GAATTCGGCTTCCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GAATTCGGCTTCCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT 60

Qy   61 ATAACTTTTATTCTACTATGTATGGAATAGTATTAATAAATGAACTAGGGAAG 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ATAACTTTTATTCTACTATGTATGGAATAGTATTAATAAATGAACTAGGGAAG 120

Qy   121 GATGTAATAAATTAGACACTTCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GATGTAATAAATTAGACACTTCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC 180

Qy   181 TTTTCTCAAATAAATATGTGAATCAAAGAAGCCAAATCAAGAGCTAAGGAAAAAGAAATGGT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TTTTCTCAAATAAATATGTGAATCAAAGAAGCCAAATCAAGAGCTAAGGAAAAAGAAATGGT 240

Qy   241 TCAGGAAAAGAGAAACAAGCAAAGAGCCAAAATCAGAGCACCGGTCCAGTAGATCTGGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TCAGGAAAAGAGAAACAAGCAAAGAGCCAAAATCAGAGCACCGGTCCAGTAGATCTGGG 300

Qy   301 CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAAGACACACCCGGTCCAGTAGATCTGGG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAAGACACACCCGGTCCAGTAGATCTGGG 360

Left column (Qy/Db alignment):

```
Qy   361  CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT  420

Qy   421  GACAAACTGCTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GACAAACTGCTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA   480

Qy   481  TTCAGTGAAGAAATATTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA   540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TTCAGTGAAGAAATATTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA   540

Qy   541  CCTCAACAAATTCACCTTAAAGCAAAGCAATATATAGAGAAATTTATACAGACTGATGCC  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  CCTCAACAAATTCACCTTAAAGCAAAGCAATATATAGAGAAATTTATACAGACTGATGCC  600

Qy   601  CCAAAAGAGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAA  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  CCAAAAGAGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAA  660

Qy   661  CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAAGACCTCAG  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAAGACCTCAG  720

Qy   721  AGTTATACACGTTTCTGAAAATCTGACATCTATTTAGACTTAGAAAATTGATTTATTC    780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  AGTTATACACGTTTCTGAAAATCTGACATCTATTTAGACTTAGAAAATTGATTTATTC    780

Qy   781  AGACCACAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA   840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  AGACCACAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA   840

Qy   841  CAATCAGATGTTGCCATTGGTTATAAAGATCGCATGTTTAATGTTAAGAAATTGGTCCCATCC  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CAATCAGATGTTGCCATTGGTTATAAAGATCGCATGTTTAATGTTAAGAAATTGGTCCCATCC  900

Qy   901  TTATACATCGCTTCTAACATATCGCATGTTTATGTTAAGAATGTAAAAAACAAAAAACTTCTGCTAACAAA  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TTATACATCGCTTCTAACATATCGCATGTTTATGTTAAGAATGTAAAAAACAAAAAACTTCTGCTAACAAA  960

Qy   961  CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAAACTTTCTGCTAACAAA  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAAACTTTCTGCTAACAAA  1020

Qy   1021 ATACACAGTATCCGCCAGTATATTCGTATAAACCTTCGTCATTGATGTCATTCCATTTA   1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ATACACAGTATCCGCCAGTATATTCGTATAAACCTTCGTCATTGATGTCATTCCATTTA   1080

Qy   1081 TAATCAGAAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAGTAATAAATGTTT  1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TAATCAGAAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAGTAATAAATGTTT  1140

Qy   1141 TATAAGAGTTGTGAGTAAGTAAGTAAAGTTAAGCTTTGCAAAGTTGTCAAACA      1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TATAAGAGTTGTGAGTAAGTAAGTAAAGTTAAGCTTTGCAAAGTTGTCAAACA      1200

Qy   1201 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAATATGTGTTATAATAAACATAATAAT  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAATATGTGTTATAATAAACATAATAAT  1260

Qy   1261 ACTCAGATATCCAAAATGTTCAGATAGCATTTTTCATAATGAATGTTCTCTTTTTTTTGGT  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ACTCAGATATCCAAAATGTTCAGATAGCATTTTTCATAATGAATGTTCTCTTTTTTTTGGT  1320

Qy   1321 AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT  1380

Qy   1381 ACTACTAACCCTGTCCCAAGAATTAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAA  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ACTACTAACCCTGTCCCAAGAATTAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAA  1440

Qy   1441 CTTTTGTGAAGACACATTCATCTCTACAGAACTTCAGATTAATAATATCTAGATTAATG  1500
```

Right column (Qy/Db alignment):

```
Db   1441 CTTTTGTGAAGACACATTCATCTCTACAGAACTTCAGATTAATAATATCTAGATTAATG  1500

Qy   1501 ACTGAGAATAGATCCACATTGACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT     1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ACTGAGAATAAGATCCACATTGACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT     1560

Qy   1561 ACAAAGTACTTCTGTTGGTCACAGAAAATGACCAGATTTGCATATCTCCAGTAGGGA    1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 ACAAAGTACTTCTGTTGGTCACAGAAAATGACCAGATTTGCATATCTCCAGTAGGGA    1620

Qy   1621 ACTAAGTAGACTACCTTATCCGGCTAAGAAAACTTGCTACTAAAACATTAGGCCATCA    1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ACTAAGTAGACTACCTTATCCGGCTAAGAAAACTTGCTACTAAAACATTAGGCCATCA    1680

Qy   1681 ATGGCTTGAATAAAAACCAGAGAAGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG   1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 ATGGCTTGAATAAAAACCAGAGAAGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG   1740

Qy   1741 AATTGGGTAGAAATCAGAAATGAGATGAGATGGTTCATATTCAGAAGTGTTATCTCATTGATATATTC  1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 AATTGGGTAGAAATCAGAAATGAGATGAGATGGTTCATATTCAGAAGTGTTATCTCATTGATATATTC  1800

Qy   1801 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAGTGTTATCTCATTGATTATATTC  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAGTGTTATCTCATTGATTATATTC  1860

Qy   1861 TTGTTAAGCAAATCCCTAAGTAATTAATTATTAATTTTAATGTGTTACATGGTCTGTAAATATT   1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TTGTTAAGCAAATCTCCTAAGTAATTAATTATTAATTTTAATGTGTTACATGGTCTGTAAATATT   1920

Qy   1921 ATGTCACTGTTTAAAGATGCCATGCATTGTAGTAATAATATTTTAACTTCATTACAGTTAAGT  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ATGTCACTGTTTAAAGATGCCATGCATTGTAGTAATAATATTTTAACTTCATTACAGTTAAGT  1980

Qy   1981 TGTATTTAAAAATGCCATGGCTTTGGAAATTAATAATATTTTAACTTCATACAGTTAAGT   2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TGTATTTAAAAATGCCATGGCTTTGGAAATTAATAATATTTTAACTTCATACAGTTAAGT   2040

Qy   2041 GTGAAAACTTTAGATCATTGTAGTAATAAAATATTTTTAACTTCATTACAGTTAAGT     2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GTGAAAACTTTAGATCATTGTAGTAATAAAATATTTTTAACTTCATTACAGTTAAGT     2100

Qy   2101 TTAATCGACAATAAAAGCTCGACTGGAAGTCTGATTATCCTTCCATTATGTAATAAGGA   2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 TTAATCGACAATAAAAGCTCGACTGGAAGTCTGATTATCCTTCCATTATGTAATAAGGA   2160

Qy   2161 ATAAACATTTTCTTCTTTTTTAGAGTAAAAAAAAAAAAAAAAAAAGGGCGGCCGC  2217
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ATAAACATTTTCTTCTTTTTTAGAGTAAAAAAAAAAAAAAAAAAAGGGCGGCCGC  2217
```

RESULT 2
US-09-498-959-1
; Sequence 1, Application US/09499959
; Patent No. 6410240
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; APPLICANT: Yowe, David
; TITLE OF INVENTION: RGS-Containing Molecules and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 5800-19A
; CURRENT APPLICATION NUMBER: US/09/498,959
; CURRENT FILING DATE: 2000-02-04
; EARLIER APPLICATION NUMBER: 09/244,314
; EARLIER FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```
; NAME/KEY: CDS
; LOCATION: (160)...(867)
US-09-498-959-1

Query Match      100.0%;   Score 2217;   DB 4;   Length 2217;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 2217;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

Qy   1 GAATTCGGCTTCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GAATTCGGCTTCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT   60

Qy   61 ATAACTTTTATTCTACTACTGTATATGATGGAATAGTTAATAAATGACTAGGGAAG   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ATAACTTTTATTCTACTACTGTATATGATGGAATAGTTAATAAATGACTAGGGAAG   120

Qy   121 GATGTAATAAATTAGACATCTCTTCATTTTAGAGGAAGAGATGGAAACAACATTGCTTTTC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GATGTAATAAATTAGACATCTCTTCATTTTAGAGGAAGAGATGGAAACAACATTGCTTTTC   180

Qy   181 TTTTCCAAATAAATATGTGAATCAAAGAAAAAAACTTTTTCAAGTTAATACATGGT   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TTTTCCAAATAAATATGTGAATCAAAGAAAAAAACTTTTTCAAGTTAATACATGGT   240

Qy   241 TCAGGAAAAGAAGAAACAAGCAAAGAGCCAAAATCAGAGCTAAGGAAAAAAAGAAATAGA   300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TCAGGAAAAGAAGAAACAAGCAAAGAGCCAAAATCAGAGCTAAGGAAAAAAAGAAATAGA   300

Qy   301 CTAAGTCTTCTTGTGCAGAAACCTGAGTTCATGAGACACCCGCTCCAGTAGATCTGGG   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTAAGTCTTCTGTGCAGAAACCTGAGTTCATGAGACACCCGCTCCAGTAGATCTGGG   360

Qy   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT   420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT   420

Qy   421 GACAAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTGGATAGCCTGTGAAGATTCTTAAAACTGAA   480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GACAAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTGGATAGCCTGTGAAGATTCTTAAAACTGAA   480

Qy   481 TTCAGTGAAGAAAATATTGAATTTGAGATTTGGATAGCCTGTGAAGATTCAGAGAAAAGCCAAGGGA   540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TTCAGTGAAGAAAATATTGAATTTGAGATTTGGATAGCCTGTGAAGATTCAGAGAAAAGCCAAGGGA   540

Qy   541 CCTCAACAAATTCACCTTAGGAGACGATCACGCATTATTATACAGACTGATGCC   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCTCAACAAATTCACCTTAGGAGACGATCACGCATTATTATACAGACTGATGCC   600

Qy   601 CCAAAAGAGGTTAACCTGATTTTCACACAAAAAGAGAGTCATTACAAACAGCATCACTCAA   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CCAAAAGAGGTTAACCTGATTTTCACACAAAAAGAGAGTCATTACAAACAGCATCACTCAA   660

Qy   661 CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAAAGCGTCATCAGCTGAACAAGAC   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAAAGCGTCATCAGCTGAACAAGAC   720

Qy   721 AGTTATACACGTTTCTGAAATCTGACATCTATTTGACTTAGACTTTAGAGTGAAG   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AGTTATACACGTTTCTGAAATCTGACATCTATTTGACTTAGACTTTAGAGTGAAG   780

Qy   781 AGACCAACAAATCTTAGGAGACACGATCACGCTCATTACCTGCAATGAATTCCAAGATGTA   840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGACCAACAAATCTCTAGGAGACACGATCACGCTCATTACCTGCAATGAATTCCAAGATGTA   840

Qy   841 CAATCAGATGTGCCCATTGTGGTATAAAGAAAATTGATTTTGCTCATTTATTGACAAAC   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CAATCAGATGTGCCCATTGTGGTATAAAGAAAATTGATTTTGCTCATTTATTGACAAAC   900

Qy   901 TTATACATTCTGCTTCTAACATATCGCATTGTTATGTTAAGATTTGTTGTCCCTATTTAAA   960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTATACATTCTGCTCTAACATATCGCATTGTTATGTTAAGATTTGGTCCCATCCTTTAAA   960

Qy   961 CTGAAAATATGTCATGTGAAATTATTTAAAAATGTAAAACAAAACTTCTGCTAACAA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CTGAAAATATGTCATGTGAAATTATTTAAAAATGTAAAACAAAACTTCTGCTAACAA 1020

Qy   1021 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCATTTGATGTCATTCCATTTA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCATTTGATGTCATTCCATTTA 1080

Qy   1081 TAATCAGAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAAAGTAATAATGTTT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TAATCAGAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAAAGTAATAATGTTT 1140

Qy   1141 TATAAGATTGTAGAGTTAAGTAAAGTTAAGCTTTTGCAAAGTGTCAAAAGTTCAAACA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TATAAGATTGTAGAGTTAAGTAAAGTTAAGCTTTTGCAAAGTGTCAAAAGTTCAAACA 1200

Qy   1201 AAAGTTCAGTTGGATTTTTACCAAAGCAGCATAATATGTGTTATAAACATAATAAT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AAAGTTCAGTTGGATTTTTACCAAAGCAGCATAATATGTGTTATAAACATAATAAT 1260

Qy   1261 ACTCAGATATCCAAATGTTCAGATAGCATTTTTCTATAATGGAACATTTATTAATTCGGGTT 1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ACTCAGATATCCAAATGTTCAGATAGCATTTTTCTATAATGGAACATTTATTAATTCGGGTT 1320

Qy   1321 AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTATTAAGCCAACAGGAA 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTATTAAGCCAACAGGAA 1380

Qy   1381 ACTACTAACCCTGTCCCAAGAATAGTAATATCACCTTCAGTTATAGAGCCAGCAACAGGAA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ACTACTAACCCTGTCCCAAGAATAGTAATATCACCTTCAGTTATAGAGCCAGCAACAGGAA 1440

Qy   1441 CTTTTGTGAAGACACATTCATCTCTACAGACTTCAGATAATAAAATAATCTAGATTAATG 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CTTTTGTGAAGACACATTCATCTCTACAGACTTCAGATAATAAAATAATCTAGATTAATG 1500

Qy   1501 ACTGAGAATAAGACTACCTTATCATTTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ACTGAGAATAAGACTACCTTATCATTTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTAT 1560

Qy   1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCCAGATTTTGCATATCTCCAGGTAGGGA 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCCAGATTTTGCATATCTCCAGGTAGGGA 1620

Qy   1621 ACTAAGTAGACTACCTTATCCGGCTAAGAAACTTGCTACTAAACTATTAGGCCATCA 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ACTAAGTAGACTACCTTATCCGGCTAAGAAACTTGCTACTAAACTATTAGGCCATCA 1680

Qy   1681 ATGGCTTGAATAAAACCAGAGAAGGTTTTCCCAGGACGTCTCATGTTGCCCTTTAG 1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 ATGGCTTGAATAAAACCAGAGAAGGTTTTCCCAGGACGTCTCATGTTGCCCTTTAG 1740

Qy   1741 AATTGGGGTAGAAATCAGAAATGAGATGAGGGAACAAGCAAGGAGTCTAGGCCTAGC 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 AATTGGGGTAGAAATCAGAAATCAGAAGGGAACAAGCAAGGAGTCTAGGCCTAGC 1800

Qy   1801 GATTTGGGCATCTGCCACATTGGTCATATTCAGAAGTGTTATCCTCATTGATTATATTC 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GATTTGGGCATCTGCCACATTGGTCATATTCAGAAGTGTTATCCTCATTGATTATATTC 1860

Qy   1861 TTGTTAAGCAAATCTCCTTAAGATATATTCAAATAAGATTATACTCATACATCTAT 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TTGTTAAGCAAATCTCTCTTAAGATATATTCAAATAAGATTATACTCATACATCTAT 1920

Qy   1921 ATGTCACTGTTTTAAGAGATATTTAATTTTTAAATGTGTTACATGGTCTGTAAATATT 1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ATGTCACTGTTTTAAGAGATATTTAATTTTTAAATGTGTTACATGGTCTGTAAATATT 1980

Qy   1981 TGTATTTAAAAATGCCATGCATGCCACATTGGTCTCCCATCCTCATTTAAA 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TGTATTTAAAAATGCCATGCATGCCACATTGGTCTCCCATCCTCATTTAAA 2040

Qy   2041 GTGAAAACTTTAGATCATTGTGTAGTAATAAAATATTTTAACTTCATCACAGTTAAGT 2100
```

```
Db  2041 GTGAAAACTTTAGATCATTTGTAGTAGTAATAAATATTTTAACTTCATTCATCATACAGTTAAGT 2100
              |||||||||||||||||||||||||||||
Qy  2101 TTATCTGACAATAAAAGCTCTGACTGAAGTTGATATCCTTCCTATTATGTAATAAGGA 2160
              |||||||||||||||||||||||||||||
Db  2101 TTATCTGACAATAAAAGTCTGACTGAAGTTGATATCCTTCCTATTATGTAATAAGGA 2160
Qy  2161 ATAAACATTTTCTCTTCTTTTAGAGTAAAAAAAAAAAAAAAAAAAAAAGGCGGGCCGC 2217
              ||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 ATAAACATTTTCTTCTTTTAGAGTAAAAAAAAAAAAAAAAAAAAAAGGCGGGCCGC 2217
```

RESULT 3
US-09-244-314-3
; Sequence 3, Application US/09244314
; Patent No. 6274362
; GENERAL INFORMATION:
;   APPLICANT: Hodge, Martin R.
;   APPLICANT: Yowe, David
;   TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
;   FILE REFERENCE: 5800-19, 035800/174680
;   CURRENT APPLICATION NUMBER: US/09/244,314
;   CURRENT FILING DATE: 1999-02-04
;   NUMBER OF SEQ ID NOS: 4
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1164
;   TYPE: DNA
;   ORGANISM: Mus sp.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (134)..(841)
US-09-244-314-3

Query Match   26.9%; Score 596; DB 3; Length 1164;
Best Local Similarity 75.8%; Pred. No. 8.8e-118;
Matches 791; Conservative 0; Mismatches 220; Indels 33; Gaps 3;

```
Qy   69 TTATTCTACTATGTATATGTGAATAGTATTAATAAATGAACTAGGGAGGGATGTAAT 128
             |||   ||  |  ||||      ||||||||||
Db   41 TTTATTCAGGATGTTTTCCTATGAAATAGCATTCATCTGTGGGAGAGAGAAGGACTAAGG 100
Qy  129 AAATTAGGACATCCTTCAT--TTTAGAGAGAAGATGGAAACACATTGCTTTTCTTTTCTT 186
             ||||  | |||| ||||    |||||||||||||||||||||||||||||
Db  101 AAATCTGACATCTGTGTTCCTGGGACAGAATATGTCACTGGT-TTCTTCTCT 160
Qy  187 CAAATAAATATGTGAATCAAAGAAAAAACTTTTTTCAGTTAATACATTGGTTCAGGA 246
             ||| |||||||||||||||||| |||||||||||||| ||| |  ||||||| ||||
Db  161 CAATTAAATATGTGAATCAAAGAGAAAACTTTTTTCAAACTAATGCATGGGTCAGGG 220
Qy  247 AAGAAGAAACAAGCAAAGAAGCCAAAATCAGAGCTAAGGAGAAAAAAGAAGACTAAGT 306
             |||||||||||||| | |||||||||||||||| |||| ||||||||||||||||||
Db  221 AAGAAGAAACAAGCATCGAGGCCAAAATCAGAGCGCAAAATCAGGAATAGACTAAGT 280
Qy  307 CTTCTTGTGCAGAAACCTGAGTTCATGAAGACACCCGCTCCAGTAGATCTGGGCACTTG 366
             ||||| | ||| |||  ||| ||| ||  ||||| |  ||||||||||||  |  |||
Db  281 CTTCTCCTACAGAGAGGCCTTGACTTCCATCGGAGAGCTCAAGCCAGTAGATCTGCCCCTCTTG 340
Qy  367 GCCAAAGAAACAAGAGTCTCCCCTGAGAGGGCAGTGAAATGGGTGAATCATTTGACAAA 426
             |||||||| |||||||||| | |||||| ||||||||||||||||||||||||||||
Db  341 GCCAAAGAACAAGAGTCTCCTGAGAAGCAGTGAAATGGGTGAATCATTTGACAAA 400
Qy  427 CTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 486
             |||| | ||||||||||||||  | | |||||||| |||||||||||||||||||||||
Db  401 TTGCTCTCTCATAGAGATGGAGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 460
Qy  487 GAAGAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGCAAAAGCAAGGGACCTCAA 546
             || |||  | ||||||||| | || |  |||||| ||||||||| | ||| |||||||
Db  461 GAGGAGAGAACATTGAATTTGGGTCGCCCTGGAAGACTTCAAGCAAAATGCAAGAAACCTCAA 520
Qy  547 CAAATTCACCTTAAAGCAAAAGCAATATAGCAAATTATACAGACTGATGCCCCAAAA 606
             |||| | ||| |||||||||| ||||| ||| |||| || |||  || ||||| ||
Db  521 CAAATCATCCTAAAAGCAAAGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCCAAA 580
```

```
Db  607 GAGGTAACCTTGATTTCCACAAAGAGTCATTACAAACAGGCATCACTCACCTACC 666
             |||||||| ||||||||  ||||||||  |||| ||| | ||||
Qy  581 GAGGTAACATTGATTTCATCATACTACAAAGAGTAATTGCTAAGACCATCGCCCAGCCCACT 640
Qy  667 CTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGACAAGACAGTTAT 726
             |||||||||||||| | | ||||||||||||||| |||||||||||||||  |||||||
Db  641 CTCCACAGTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGACAATGACAGTTAT 700
Qy  727 ACACGTTTTCTGAAATCTGACATCTTATTTAGACTTGATGGAAGGAAGACCTCAGAGACCA 786
             | ||| |||| |||||||||  ||| | ||| | ||| | ||||||||||||||||||||||
Db  701 AAACGCGTTTTGAAATCTGAAATCTGAGACCTACATTGATGAGAGGAAGACCTCAGAGACCA 760
Qy  787 ACAAATCTTAGGAGACGATCACGCTCCATTACCTGCAATGAATTCCAAGATGTACAATCA 846
             |||| ||||||||||||||||| |||| ||| ||| || |||||||||||||  | |||
Db  761 ACAAACCTTAGGAGACGATCACGATCATTTACTTACTACATTCCAAGATGTAAAGTCA 820
Qy  847 GATGTTGCCATTGGTTAATAAAGAAAATTGATTTTGCTCATTTTTATGACAAACTTATAC 906
             |||||||||||| ||||| || | |||| ||||||| |||| | ||| |  |||| ||||
Db  821 GATGTTGCCATTTGGTTATGATGTAAAAGTCATTTTGTCTTCTTTTTGATAGTGTATGTGTAT 880
Qy  907 ATC----------------------------TGCTTCTAACATATCCGCATGTTTATGTTAAGA 941
             |||                            |||||||| ||| | |  ||| || |||||
Db  881 ATCTAAAATATATACTAATACTAATGTGTACTTCTAAAATATAGCTTGTGTATAAGAAGA 940
Qy  942 TTTGGTCCCATCCTTTAAACTGAAATGTCATGTGAAATTATTTTAAAAATGTAAAAAC 1001
             |||  | ||| ||| ||||||||  |||| |||| || | ||||  ||||||| || ||
Db  941 GATGATTTCATTTT--------TAAAATACACCATGCAAATACATATTAATGTAAGAAACTT 994
Qy  1002 AAAACTTTCTGCTAACAAAATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTA 1061
              |||| |||| ||| ||| | || | | | |||| | ||| ||||| ||| ||| ||| |
Db  995 TTTATTATTACTAACATAAAATAATTCATCATCCTATCTTCCGAAATATTTATGAAAATCTATC 1054
Qy  1062 TTTGATGTCATTCCATTTATAATC 1085
             |||||||||||| || |||||||
Db  1055 TGATATTCATTCTAATAAAATTC 1078
```

RESULT 4
US-09-498-959-3
; Sequence 3, Application US/09498959
; Patent No. 6410240
; GENERAL INFORMATION:
;   APPLICANT: Hodge, Martin R.
;   APPLICANT: Yowe, David
;   TITLE OF INVENTION: RGS-Containing Molecules and Uses
;   TITLE OF INVENTION: Thereof
;   FILE REFERENCE: 5800-19A
;   CURRENT APPLICATION NUMBER: US/09/498,959
;   CURRENT FILING DATE: 2000-02-04
;   EARLIER APPLICATION NUMBER: 09/244,314
;   EARLIER FILING DATE: 1999-02-04
;   NUMBER OF SEQ ID NOS: 12
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 1164
;   TYPE: DNA
;   ORGANISM: Mus sp.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (134)...(841)
US-09-498-959-3

Query Match   26.9%; Score 596; DB 4; Length 1164;
Best Local Similarity 75.8%; Pred. No. 8.8e-118;
Matches 791; Conservative 0; Mismatches 220; Indels 33; Gaps 3;

```
Qy   69 TTATTCTACTATGTATATGTGAATAGTATTAATAAATGAACTAGGGAGGGATGTAAT 128
             |||   ||  |  ||||      ||||||||||
Db   41 TTTATTCAGGATGTTTTCCTATGAAATAGCATTCATCTGTGGGAGAGAGAAGGACTAAGG 100
Qy  129 AAATTAGGACATCCTTCAT--TTTAGAGAGAAGATGGAAACAACATTGCTTTTCTTTTCT 186
             ||||  | |||| ||||    |||||||||||||||||||||||||||||
Db  101 AAATCTGACATCTGTGTTGGTGTTCCTATGAAATAGCATTCATCTGTGGGAGAAATAGCTATC 160
```

; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1010
;   LENGTH: 736
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(736)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1010

Query Match   17.5%; Score 388; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.2e-73;
Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    481 TTCAGTGAAGAAAATATTGAATTTGATAGCCTGTGAAGATTCAAGAAAAGCAGGGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4 TTCAGTGAAGAAAATATTGAATTTGATAGCCTGTGAAGATTCAAGAAAAGCAAGGGA 63

Qy    541 CCTCAACAAATTCACCTAAAGCAAAAGCAATAATATGAGAAATTTATACAGACTGATGCC 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 CCTCAACAAATTCACCTAAAGCAAAAGCAATAATATGAGAAATTTATACAGACTGATGCC 123

Qy    601 CCAAAAGAGGTAACCTGATTTCACACAAAGAAGTCATTACAAACAGCATCACTCAA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 CCAAAAGAGGTAACCTGATTTCACACAAAGAAGTCATTACAAACAGCATCACTCAA 183

Qy    661 CCTACCCTCCACAGTTTGATGCTGCACAAAGCAGCAGTGTATCAGCTCATGGAACAAGAC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 CCTACCCTCCACAGTTTGATGCTGCACAAAGCAGCAGTGTATCAGCTCATGGAACAAGAC 243

Qy    721 AGTTATACACGTTTCTGAAATCTGACACTATTTAGACTTGATGGAGGAGGAGACCTCAG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 AGTTATACACGTTTCTGAAATCTGACATCATTTAGACTTGATGGATGGAGGAGACCTCAG 303

Qy    781 AGACCACAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 AGACCACAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA 363

Qy    841 CAATCAGATGTTGCCATTTGGTTATAAA 868
          ||||||||||||||||||||||||||||
Db    364 CAATCAGATGTTGCCATTTGGTTATAAA 391
```

RESULT 6
US-09-016-434-1278
; Sequence 1278, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
;   APPLICANT: Janice Au-Young
;   APPLICANT: Jeffrey J. Seilhamer
;   TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
;   TITLE OF INVENTION: PATHWAY GENE EXPRESSION
;   NUMBER OF SEQUENCES: 1490
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;   CURRENT APPLICATION DATA:

```
Qy    187 CAAATAATATGTGTGAATCAAAAGAAAAACTTTTTTCAAGTTAATACACATGGTTCAGGA 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 CAATTAAATATGTGTGAATCAAAAGAAAAACTTTTTTCAAACTAATACATGGGTCAGGG 220

Qy    247 AAAGAAGAAACAAGCAAAGAAGCCAAAATCAGAGCTAAGGAAAAAGAAAATAGACTAAGT 306
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    221 AAAGAGAGAAACAAGCATCGAGCGCAAAGAGGCAAAGAAGCGAAAGAAAATAGAACTAAGT 280

Qy    307 CTTCTTGTGCAGAAAACTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGGCACTTG 366
          |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281 CTTCTCCCACAGAGAGGGCCTGACTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCTCTTG 340

Qy    367 GCCAAAGAAACAAGAGTCTCCCTGAAGAGGCAGTGAATGGGGTGAATCATTGACAAA 426
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341 GCCAAAGAAACAAGAGTCTCCTGAAGAAGCAGTGAAATGGGCTGAAATCATTTGACAAA 400

Qy    427 CTGCTTTCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 486
          ||||||||||||||||||||||||||||||||  |||||||||||||||||||||
Db    401 TTGCTGTCTCTCATAGAGATGGGAGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 460

Qy    487 GAAGAAATATTGAATTTGATAGCCGTGAAGATTCAAGAAAAGCAAGGGACCTCAA 546
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    461 GAGGAGAACATTGAATTTGGGTGCCCTGAGAACTTCAAGAAATGCAAGGAACCTCAA 520

Qy    547 CAAATTCACCTAAAGCAAAAGCAATAATATGAGAAATTTATACAGACTGATGCCCCCAAAA 606
          ||||||| |||||||||||||||||||||||||||||||||||||||||| ||||||
Db    521 CAAATCATCCTAAAGCAAAAGGCAATCTCATGCAGAATTCAGGATGCTGCCCCCAAA 580

Qy    607 GAGGTAACCTGATTTCACACAAAGAAGTCATTACAAACAGCATCACTCAACCTACC 666
          |||||||||||||||||||||||||||||||| |||||||| |||||||||||
Db    581 GAGGTTAACATTGATTTTCATACTCAAAGAGTAATTGCTAAGAAGTAATTGCTCAAGTGCCCCACT 640

Qy    667 CTCCACAGTTTGATGCTGCACAAAGCAGTGTATCAGCTCATGGAACACAGTTAT 726
          ||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    641 CTCCACAGTTTGATACGGCCACACAAAGCAGTGTATCAGCCACATGACAGTTAT 700

Qy    727 ACACGTTTTCTGAAATCTGACATCTATTGATGGAGGAAGACCTCAGAGACCA 786
          |||||||||||||||||||||| |||||||||||||||||||||||||
Db    701 AAACGCTTTTTGAAATCTGAGAACTCTACTTACATTGATAGGAGAAGACCTCAGAGACCA 760

Qy    787 ACAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTACAATCA 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    761 ACAAACCTTAGGAGACGACGATCACGCTCATTTACTACTTTACCTGAATTACCTGCAATGTAAGTACA 820

Qy    847 GATGTTGCCATTTGGTTATAAAGAAAATGATTTGCTCATTTTATGACAAACTTATAC 906
          |||||||||||||||||||||||||||||| ||||| ||||| |||
Db    821 GATGTTGCCATTTGGTTATGAGTAAAAGTCATTTGTCTTCTTTTGTAGTGTATGTGTAT 880

Qy    907 ATC------------------------------------TGCTTCTAACATATCGCATGTTTATGTTAAGA 941
          |||                                    ||| ||||
Db    881 ATCTAAAAATATATACTAATACTAATGTGTACTTCCTAAAATAATATAGCTTTGTGTATAAGAAGA 940

Qy    942 TTTGGGTCCCATCCTTTAAACTGAAATATGTCATGTGAAAATTATTTAAAAATGTAAAAAC 1001
          ||||||| |||||||||||| ||||||||| |||||||||||||||||||||||||||||
Db    941 GATGAATTTCATTT------TAAAATACACCATGCAAATACATATTAAATGTAAGAACTT 994

Qy   1002 AAAACTTTTCTGCTAACAAAATACATACAGTATGCCAGTATATTCTGTAAAACCTTCTA 1061
          ||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    995 TTTATATTATACTAACAAAATAATTCATCATCTATCTTCCGAAATATTTTTATGAAAATCTATC 1054

Qy   1062 TTTGATGTCATTCATTTATAATC 1085
          |||||||||||||||||||||
Db   1055 TGATATTCATTCTAATAAAAATTC 1078
```

RESULT 5
US-09-833-381-1010
; Sequence 1010, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;   APPLICANT: Robison, Keith E.
;   TITLE OF INVENTION: No. 6672186e1 Nucleic Acid and Protein Homologs

STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/023,655
  FILING DATE: HEREWITH
  CLASSIFICATION:
PRIOR APPLICATION DATA:
  APPLICATION NUMBER:
  FILING DATE:
  CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
  NAME: Zeller, Karen J.
  REGISTRATION NUMBER: 37,071
  REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (650) 855-0555
  TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1243:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1345 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  IMMEDIATE SOURCE:
    LIBRARY: GENBANK
    CLONE: g292054
US-09-023-655-1243

Query Match   7.0%;   Score 155.6; DB 4; Length 1345;
Best Local Similarity   63.2%;   Pred. No. 3.3e-24;
Matches 239; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

Qy   385 TCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTTGACAAACTGCTTTCCATAGAGAT 444
          |||||| ||| || ||| || | ||| ||| | | |||||| | ||| | |||||
Db   249 TCTCCTGAGGAAGCACAGCTGTGTGCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy   445 GGACTAGAGGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAAGAAAATATTGAATTT 504
          || | || ||| || || |||| ||| ||| | | || ||| || |||||| |||||
Db   309 GGTCTTGCTGCATTCAGGGCTTTTTAAAGTCGGAATTCTGTGAAGAAAAATATTGAATTC 368

Qy   505 TGGATAGCCTGTGTGAAGATTCAAGAAAAGCAAGGACCTCAACAAATTCACCTTAAAGCA 564
          || | ||| ||| |||||| ||| || || || | | | ||||| |||| |||||
Db   369 TGGCTGGCCTGTGAAGACTTCAAGAACTCAAAAAACCAAAAAACCAAAAAAGCTGTCCTCAAAAGCA 428

Qy   565 AAAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAAGAGGTTAACCTTGATTTT 624
          ||| |||||| ||| ||| | ||| | ||||| | ||||| || || || |||||
Db   429 AGGAAAATATATACTGACTTCATAGAAAAGGAAGCTCCAAGAAGATAAACATAGATTTT 488

Qy   625 CACACAAAGAGAGTCATTACAAACAGCATCACTCCACGTTTTGATGCT 684
          |||| ||| ||| | ||| ||| ||| | || ||| || |||||||
Db   489 CAAACCAAAACTCTGATTGCCCAGAATATACAAGGTGGCTGCTTACAACT 548

Qy   685 GCACAAAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTATACACGTTTTCTGAAATCT 744
          ||||| || ||| | ||| || | |||| |||||| ||||| ||| || ||||| |||
Db   549 GCCCAGAAAAGGGTATACAGCTTGATGAGAACAACTCTTATCCTCGTTTCTTGGAGTCA 608

Qy   745 GACATCTATTTAGACTTG 762
          || | | || ||||||
Db   609 GAATTCTTACCAGGACTTG 626

RESULT 8
US-08-748-483-2
; Sequence 2, Application US/08748483
; Patent No. 5955314
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.

APPLICATION NUMBER: US/09/016,434
  FILING DATE: HEREWITH
  CLASSIFICATION:
PRIOR APPLICATION DATA:
  APPLICATION NUMBER:
  FILING DATE:
  CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
  NAME: Zeller, Karen J.
  REGISTRATION NUMBER: 37,071
  REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (650) 855-0555
  TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1278:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1345 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  IMMEDIATE SOURCE:
    LIBRARY: GENBANK
    CLONE: g292054
US-09-016-434-1278

Query Match   7.0%;   Score 155.6; DB 4; Length 1345;
Best Local Similarity   63.2%;   Pred. No. 3.3e-24;
Matches 239; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

Qy   385 TCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTTGACAAACTGCTTTCCATAGAGAT 444
          |||||| ||| || ||| || | ||| ||| | | |||||| | ||| | |||||
Db   249 TCTCCTGAGGAAGCACAGCTGTGTGCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy   445 GGACTAGAGGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAAGAAAATATTGAATTT 504
          || | || ||| || || |||| ||| ||| | | || ||| || |||||| |||||
Db   309 GGTCTTGCTGCATTCAGGGCTTTTTAAAGTCGGAATTCTGTGAAGAAAAATATTGAATTC 368

Qy   505 TGGATAGCCTGTGTGAAGATTCAAGAAAAGCAAGGACCTCAACAAATTCACCTTAAAGCA 564
          || | ||| ||| |||||| ||| || || || | | | ||||| |||| |||||
Db   369 TGGCTGGCCTGTGAAGACTTCAAGAACTCAAAAAACCAAAAAACCAAAAAAGCTGTCCTCAAAAGCA 428

Qy   565 AAAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAAGAGGTTAACCTTGATTTT 624
          ||| |||||| ||| ||| | ||| | ||||| | ||||| || || || |||||
Db   429 AGGAAAATATATACTGACTTCATAGAAAAGGAAGCTCCAAGAAGATAAACATAGATTTT 488

Qy   625 CACACAAAGAGAGTCATTACAAACAGCATCACTCCACGTTTTGATGCT 684
          |||| ||| ||| | ||| ||| ||| | || ||| || |||||||
Db   489 CAAACCAAAACTCTGATTGCCCAGAATATACAAGGTGGCTGCTTACAACT 548

Qy   685 GCACAAAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTATACACGTTTTCTGAAATCT 744
          ||||| || ||| | ||| || | |||| |||||| ||||| ||| || ||||| |||
Db   549 GCCCAGAAAAGGGTATACAGCTTGATGAGAACAACTCTTATCCTCGTTTCTTGGAGTCA 608

Qy   745 GACATCTATTTAGACTTG 762
          || | | || ||||||
Db   609 GAATTCTTACCAGGACTTG 626

RESULT 7
US-09-023-655-1243
; Sequence 1243, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO

APPLICANT: Goli, Surya K.
TITLE OF INVENTION: A NOVEL REGULATOR OF CELL SIGNALLING
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
  ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
  STREET: 3174 Porter Drive
  CITY: Palo Alto
  STATE: CA
  COUNTRY: US
  ZIP: 94304
COMPUTER READABLE FORM:
  MEDIUM TYPE: Diskette
  COMPUTER: IBM Compatible
  OPERATING SYSTEM: DOS
  SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/748,483
  FILING DATE: Herewith
  CLASSIFICATION: 530
PRIOR APPLICATION DATA:
  APPLICATION NUMBER:
  FILING DATE:
ATTORNEY/AGENT INFORMATION:
  NAME: Billings, Lucy J.
  REGISTRATION NUMBER: 36,749
  REFERENCE/DOCKET NUMBER: PF-0157 US
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415-855-0555
  TELEFAX: 415-845-4166
  TELEX:
IMMEDIATE SOURCE:
  LIBRARY: Consensus
  CLONE: 57362
US-08-748-483-2

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
  LENGTH: 744 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear

Query Match   6.9%; Score 153.2; DB 2; Length 744;
Best Local Similarity 61.6%; Pred. No. 9.3e-24;
Matches 245; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

```
Qy  367 GCCAAAGAAACAAGAGTCTCCCTGAAGAGGCAGTGAATGGGTGAATCATTTGACAAA 426
            ||||  ||  ||||||   ||| || ||||||  ||| |   ||   |||| ||||||
Db  210 GCCAAGACCCAGAAAACCTCGCTGGACGAGGCCCTGCAGTGGCGTGATTCCCTGGACAAA 269

Qy  427 CTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 486
            || |   || ||| | ||||||| ||| ||  ||| ||| ||  | |||| ||||||||||
Db  270 CTCCTGCAGAACAACTATGGACTTGCCAGTTTCAAAAGTTTCCTGAAGTCTGAATTCAGT 329

Qy  487 GAAGAAATATTGAATTTGGATAGCCTGTGAAGATTCAAGAAAAGCAAGGGACCTCAA 546
            || ||||  || ||| |||||| |||||| |||| |||||  || ||||   ||| ||
Db  330 GAGGAAAAACCTTGAGTTCTGGATTGCCTGTGAGGATTACAAGAAGAAGATCAAGTCCCCTGCC 389

Qy  547 CAAATTCACCTTAAAGCAAAAGCAATATATGACGAATTTATACAGACTGATGCCCCAAAA 606
            | |||||| ||||||  |||| ||| ||| ||  ||  || |||| | | ||  ||| ||
Db  390 AAGATGGCTGAGAAGGCAAAGCAAATTTATGAAGAATTCATTCAAACAGCTCCTAAA 449

Qy  607 GAGGTTAACCTTGATTTCACACAAAAGAAGTCATTCAAACAACAGCATCACTCAACCTACC 666
            || || |||| ||| || |||  || || || | |  |||||| |||| | | |||| | 
Db  450 GAGGTGAATATTGACCACTTCACTAGGAGACATCACAATGAAGAGAACCTGGTGGAACTTCC 509

Qy  667 CTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTAT 726
            || |||| |||| |  | |||  || || |||  ||| ||  || |||| || ||  | 
Db  510 CTGAGCAGCTTGACATGGCCCAGAAAGAATCCATGCCCTGATGGAAAAAGGATTCTG 569

Qy  727 ACACGTTTTCTGAAATCTGACATCTATTTAGACTTGAT 764
            |||  |   |  || | | |||  |  ||| || |||| 
Db  570 CCTCGCTTTGTGCGGCTCTGAGTTTTATCAGGAGTTAAT 607
```

RESULT 9
US-09-023-655-795
; Sequence 795, Application US/09023655
; Patent No. 6607879
GENERAL INFORMATION:
  APPLICANT: Cocks, Benjamin G.
  APPLICANT: Susan G. Stuart
  APPLICANT: Jeffrey J. Seilhamer
  TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
  TITLE OF INVENTION: EXPRESSION
  NUMBER OF SEQUENCES: 1508
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
    STREET: 3174 PORTER DRIVE
    CITY: PALO ALTO
    STATE: CALIFORNIA
    COUNTRY: USA
    ZIP: 94304
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/023,655
    FILING DATE: HEREWITH
    CLASSIFICATION:
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:
    FILING DATE:
    CLASSIFICATION:
  ATTORNEY/AGENT INFORMATION:
    NAME: Zeller, Karen J.
    REGISTRATION NUMBER: 37,071
    REFERENCE/DOCKET NUMBER: PA-0001 US
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (650) 855-0555
    TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 795:
SEQUENCE CHARACTERISTICS:
  LENGTH: 2190 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
IMMEDIATE SOURCE:
  LIBRARY: SYNORAT03
  CLONE: 696878
US-09-023-655-795

Query Match   6.9%; Score 153.2; DB 4; Length 2190;
Best Local Similarity 61.6%; Pred. No. 1.2e-23;
Matches 245; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

```
Qy  367 GCCAAAGAAACAAGAGTCTCCCTGAAGAGGCAGTGAATGGGTGAATCATTTGACAAA 426
            ||||  ||  ||||||   ||| || ||||||  ||| |   ||   |||| ||||||
Db  210 GCCAAGACCCAGAAAACCTCGCTGGACGAGGCCCTGCAGTGGCGTGATTCCCTGGACAAA 269

Qy  427 CTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 486
            || |   || ||| | ||||||| ||| ||  ||| ||| ||  | |||| ||||||||||
Db  270 CTCCTGCAGAACAACTATGGACTTGCCAGTTTCAAAAGTTTCCTGAAGTCTGAATTCAGT 329

Qy  487 GAAGAAATATTGAATTTGGATAGCCTGTGAAGATTCAAGAAAAGCAAGGGACCTCAA 546
            || ||||  || ||| |||||| |||||| |||| |||||  || ||||   ||| ||
Db  330 GAGGAAAAACCTTGAGTTCTGGATTGCCTGTGAGGATTACAAGAAGAAGATCAAGTCCCCTGCC 389

Qy  547 CAAATTCACCTTAAAGCAAAAGCAATATATGACGAATTTATACAGACTGATGCCCCAAAA 606
            | |||||| ||||||  |||| ||| ||| ||  ||  || |||| | | ||  ||| ||
Db  390 AAGATGGCTGAGAAGGCAAAGCAAATTTATGAAGAATTCATTCAAACAGCTCCTAAA 449

Qy  607 GAGGTTAACCTTGATTTCACACAAAAGAAGTCATTCAAACAACAGCATCACTCAACCTACC 666
            || || |||| ||| || |||  || || || | |  |||||| |||| | | |||| | 
Db  450 GAGGTGAATATTGACCACTTCACTAGGAGACATCACAATGAAGAGAACCTGGTGGAACTTCC 509
```

Qy   667 CTCCCAAGTGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGAACAACAGACAGTTAT 726
               ||| ||||| |||||
Db   510 CTGAGCAGCTTTGACATGGCCCAGAAAAGAATCCATGCCCTGAATGGAAAAGGATTCTCTG 569

Qy   727 ACACGTTTTCTGAAATCTGACATCTATTTAGAGACTTGAT 764
              ||||||| ||||
Db   570 CCTCGCTTGTGCGCTCTGAGTTTTATCAGGAGGTTAAT 607

RESULT 10
US-09-566-921-111
; Sequence 111, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
APPLICANT: Loring, Jeanne F.
APPLICANT: Tingley, Debora W.
APPLICANT: Edwards, Carla M.
TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
FILE REFERENCE: PA-0024 US
CURRENT APPLICATION NUMBER: US/09/566,921
CURRENT FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PERL Program
SEQ ID NO 111
LENGTH: 3208
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6682888 900146.4
US-09-566-921-111

Query Match   6.6%; Score 146.6; DB 4; Length 3208;
Best Local Similarity 62.3%; Pred. No. 3.4e-22;
Matches 230; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

Qy   396 GGCAGTGAAATGGGGTGAATCATTTGACAAACTGCTTTCCCATAGAGATGGACTAGAGGC 455
              |||||||||||||||||| ||||
Db   320 GGCCCTGCAGTGGCGTGATTCCCTGGACAAACTCCTGGACAACAACTATGGACTTGCCAG 379

Qy   456 TTTTACCAGATTTCTTAAAACTGAATCAGTGAGAAAAATATTGAATTTGGATAGCCTG 515
Db   380 TTTCAAAAGTTTCTGAAGTCTCAGTGAGAGAAAACCTGAATTCTGAGTTGCCTG 439

Qy   516 TGAAGAATTCAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAAAGCAAAGCAATATA 575
Db   440 TGAGGATTACAAGAGAATCAAGTCCCCTGCCAAGATGCTGAGAGGCAAAGCAATTTA 499

Qy   576 TGAGAAATTTATACAGACTGATTTCACACAAAGA 635
Db   500 TGAAGAATTCATTCAAACGGAGGTCCTAAAGAGGTGAATATTGACCACTTCACTAAGGA 559

Qy   636 AGTCATTACAAACAGCATCACTCAACCTACCCTCCACAGTTTGATGCTGCACAAAGCAG 695
Db   560 CATCACAATGACAAGCGTGTGAACCCTGGTGAACCTTCCCTGAACAGCTTTGACAGCGCCCAGAAAAG 619

Qy   696 AGTGTATCAGCTCATGGAACAACAGCAGTTATACACGTTTTTCTGAAATCTGACATCTATTT 755
Db   620 AATCCATGCCCTGATGGAAAAGGATTCTCTGATGCGCTCTGAGTTTTATCA 679

Qy   756 AGACTTGAT 764
Db   680 GGAGTTAAT 688

RESULT 11
US-09-016-434-1091
; Sequence 1091, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1091:
SEQUENCE CHARACTERISTICS:
LENGTH: 800 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1216372
US-09-016-434-1091

Query Match   6.5%; Score 143.8; DB 4; Length 800;
Best Local Similarity 57.2%; Pred. No. 9.4e-22;
Matches 281; Conservative 0; Mismatches 207; Indels 3; Gaps 1;

Qy   277 AGAGCTAAGGAAAAAAGAAATAGACTAAGTCTTCTGTGCAGAAACCTGAGTTTCATGAA 336
Db   140 AGTGCAAAAGATATGAAACATCGGCTAGGTTTCCTGCTGCAAAAATCTGATTCCTGTGAA 199

Qy   337 GACA--CCCGCTCCAGTAGATCTGGGCACTTGCCAAAGAAACAAGAGAGTCTCCCCTGAA 393
Db   200 CACAATTCTTCCCACAACAAGAACAAGTGGTTATTTGCCGAGAGTGAGCCAAGAG 259

Qy   394 GAGGCAGTGAAATGGGGTGAATCATTTGACAAACTGCTTTCCCATAGAGATGGACTAGAG 453
Db   260 GAAGTCAAGAAATGGGCGTGAATCACTGGAAAACCTGGAATTGAATTTGGATAGCC 319

Qy   454 GCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGAAAATATTGAATTTGGATAGCC 513
Db   320 GCTTTCAAAAGTTTCTTGAAGTCTGAATATAGTGACGAGAATATTGACTTCTGGATCAGC 379

Qy   514 TGTGAAGAATTCAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAAAGCAAAGCAATA 573
Db   380 TGTGAAGAGTACAAGAGAATCAAAATCAAATCACCATCTAAACTAAGTCCCAAGGCAAAAGATC 439

Qy   574 TATGAGAAATTTATACAGACTGATGCGCCCAGGCAACCAAAGAGGTAACCTTAACCTTGA 633
Db   440 TATAATGAATTCATCTCAGTCCAGCAACAACCAAGAGGTGAACCTGGATTCTTGACCAGG 499

Qy   634 GAAGTCATTACAAACAGCATCAACCTACCCTCCACAGTTTGATGCTGCACAAAGC 693
Db   500 GAAGAGACAAGCCGAACATGCTAGAGCCTACAATAACCTGCTTTGATGAGGCCCAGAAG 559

Qy   694 AGAGTGTATCAGCTCATGGAACAAGACAGTTATCACGTTTTCTGAAATCGACATCTAT 753
           |||||||||||||||||||||| |||||||| ||||| |||||||||||| |||||
Db   560 AAGAGTTTTCAACCTGATGGAGAAGAATTCCTACCGCCGCTTCCTCAAGTCTCGATTCTAT 619

Qy   754 TTAGACTTGAT 764
           ||||||||||
Db   620 CTTGATTTGGT 630

RESULT 12
US-09-016-434-1290
; Sequence 1290, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/016,434
;   FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1290:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1398 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   LIBRARY: GENBANK
;   CLONE: G299704
US-09-016-434-1290

Query Match   6.4%;   Score 141.2; DB 4;   Length 1398;
Best Local Similarity 62.7%;   Pred. No. 3.8e-21;
Matches 237;   Conservative 0;   Mismatches 138;   Indels 3;   Gaps 1;

Qy   381 AGTCTCCCTGAAGAGGCAGTGAAATGGGTGAATCATTGACAAACTGCTTTCCCATAG 440
           ||||||||| |||||||||||||||||||||| |||||||| |||||| ||||||
Db   278 ACTTTCTGCTCGAAGTAATGCAATGGTCTCAATCTCTGGAAAAACTTCTTGCCAACCA 337

Qy   441 AGATGGACTAGAGCTTTACCCAGATTTCTTAAAACTGAATTCAGTGAAGAAAATATTGA 500
           || ||||||||||||| ||||||||||||||||||||| |||||||||||||||||||
Db   338 AACTGGTCAAAATGTCTTTGGAAGTTTCCTAAAGTCTAAAGTCAGTGAGGAGAATATTGA 397

Qy   501 ATTTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAA 560
           ||||||||||||||||||||| ||||| |||||||||||||||| ||||||||||||
Db   398 GTTCTGCTGGCTTGTGTGAAGACTATAAGAAAAACAGAG---TCTGATCTTTTGCCCTGTAA 454

Qy   561 AGCAAAAGCAATATATGAGAAATTATACGACTGATGCCCCAAAAGAGGTTAACCTTGA 620
           ||||||||||||||||||||||||||||| |||||||||||||| |||||||||||||
Db   455 AGCAGAAGAGATATATAAAGCATTTGTGCATTCAGATGCTGCTGCTCAAATCAATATTGA 514

Qy   621 TTTTCACCAAAAGAAGTCATTACAGATCAGCATCACAAACAGCAGAAGAGTTTGCATTA 680
           ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   515 CTTCCGCACTCGAGAATCTGAGAATCTTGTTATATCTTGAAGACCACCCCCACGTGTTTGA 574

Qy   681 TGCTGCACAAAGCAGCAGTGTATCAGCTCATGGAACAAGACAGTTATCACGTTTTCTGAA 740
           | ||||| |||||||||||| ||| |||||||| ||||| |||||||||||||||
Db   575 TGAAGCACAAAAGTCATCATATATACTCTTATGTGAAAAAGGACTCTTATCCCAGGTTCCTCAA 634

Qy   741 ATCTGACATCTATTAGA 758
           |||||||||||||||||
Db   635 ATCAGATATTTACTTAAA 652

RESULT 13
US-09-702-705-322
; Sequence 322, Application US/09702705
; Patent No. 654010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 322
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-322

Query Match   6.4%;   Score 141.2; DB 4;   Length 1398;
Best Local Similarity 62.7%;   Pred. No. 3.8e-21;
Matches 237;   Conservative 0;   Mismatches 138;   Indels 3;   Gaps 1;

Qy   381 AGTCTCCCTGAAGAGGCAGTGAAATGGGTGAATCATTGACAAACTGCTTTCCCATAG 440
           ||||||||| |||||||||||||||||||||| |||||||| |||||| ||||||
Db   278 ACTTTCTGCTCGAAGTAATGCAATGGTCTCAATCTCTGGAAAAACTTCTTGCCAACCA 337

Qy   441 AGATGGACTAGAGCTTTACCCAGATTTCTTAAAACTGAATTCAGTGAAGAAAATATTGA 500
           || ||||||||||||| ||||||||||||||||||||| |||||||||||||||||||
Db   338 AACTGGTCAAAATGTCTTTGGAAGTTTCCTAAAGTCTAAAGTCAGTGAGGAGAATATTGA 397

Qy   501 ATTTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAA 560
           ||||||||||||||||||||| ||||| |||||||||||||||| ||||||||||||
Db   398 GTTCTGCTGGCTTGTGTGAAGACTATAAGACTATAAGAAAAACAGAG---TCTGATCTTTTGCCCTGTAA 454

Qy   561 AGCAAAAGCAATATATGAGAAATTATACGACTGATGCCCCAAAAGAGGTTAACCTTGA 620
           ||||||||||||||||||||||||||||| |||||||||||||| |||||||||||||
Db   455 AGCAGAAGAGATATATAAAGCATTTGTGCATTCAGATGCTGCTGCTCAAATCAATATTGA 514

Qy   621 TTTTCACCAAAAGAAGTCATTACAGATCAGCATCACAAACAGCAGAAGAGTTTGCATTA 680
           ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   515 CTTCCGCACTCGAGAATCTGAGAATCTTGTTATATCTTGAAGACCACCCCCACGTGTTTGA 574

Qy   681 TGCTGCACAAAGCAGCAGTGTATCAGCTCATGGAACAAGACAGTTATCACGTTTTCTGAA 740
           | ||||| |||||||||||| ||| |||||||| ||||| |||||||||||||||
Db   575 TGAAGCACAAAAGTCATCATATATACTCTTATGTGAAAAAGGACTCTTATCCCAGGTTCCTCAA 634

Qy   741 ATCTGACATCTATTAGA 758

Db   635 ATCAGATATTTACTTAAA 652
              ||| || ||||| ||| ||||

RESULT 14
US-09-736-457-322
; Sequence 322, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Fanger, Gary
;   APPLICANT: Vedvick, Tom
;   APPLICANT: Carter, Darrick
;   APPLICANT: Retter, Marc
;   APPLICANT: Mannion, Jane
;   APPLICANT: Fan, Liqun
;   APPLICANT: Wang, Aijun
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;   TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C15
;   CURRENT APPLICATION NUMBER: US/09/736,457
;   CURRENT FILING DATE: 2000-12-13
;   NUMBER OF SEQ ID NOS: 1864
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 322
;   LENGTH: 1398
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-736-457-322

Query Match   6.4%;   Score 141.2; DB 4; Length 1398;
Best Local Similarity 62.7%;   Pred. No. 3.8e-21;
Matches 237; Conservative   0; Mismatches 138; Indels   3; Gaps   1;

Qy   381 AGTCTCCCTGAAGAGGCAGTGAATGGGGTGAATCATTTGACAAACTGCTTTCCCATAG 440
              ||| ||| |||||||| |||| |||||| |||| || ||| | ||||
Db   278 ACTTTCTGCTGCTGAGTAATGCAATGGTCTCAATCTCTGGAAAAACTTCTTGCCAACCA 337

Qy   441 AGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAAGAAATATTGA 500
              |||| ||| ||| ||| ||| |||| ||| || ||| |||||||||||||
Db   338 AACTGGTCAAAATGTCTTTGGAAGTTTCCTAAAGTCTGAATTCAGTGAGGAGAATATTGA 397

Qy   501 ATTTTGGATAGCCTGTGAAGATTCAAGAAAAGCAAGGACCTCAACAAATTCACCTTAA 560
              |||||| || |||| ||| |||| ||||||| ||| |||||| ||||
Db   398 GTTCTGGCTGGCTGTGTGAAGACTATAAGAAAACAGAG--TCTGATCTTTTGCCCTGTAA 454

Qy   561 AGCAAAAGCAATATGATGAGAAATTTATACAGACTGATGCCCCAAAAGAGGTTAACCTTGA 620
           |||||| ||||| ||| ||| | ||| |||| || ||||| || || |||||
Db   455 AGCAGAAGAGATATATAAAGCATTTGTGCATTCAGATGCTGCTAAACAAATCAATATTGA 514

Qy   621 TTTTCACACAAAGAAGTCATTACAAACAGCATCCACCTCACCTTTGA 680
           |||| ||| |||| |||| || |||||| ||| ||||
Db   515 CTTCCGCACTCGAGAATCTCAGCCAAGAAGATTAAAGCACCACCCCCAGTGTTTTGA 574

Qy   681 TGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTATACACGTTTTCTGAA 740
           ||| |||||| || ||| || ||| ||||| ||| ||| || ||||
Db   575 TGAAGCACAAAAGTCATATATACTCTTATGGAAAAGGACTCTTATCCCAGGTTCCTCAA 634

Qy   741 ATCTGACATCTATTTAGA 758
           ||| || ||||| ||| ||||
Db   635 ATCAGATATTTACTTAAA 652

RESULT 15
US-09-614-124B-322
; Sequence 322, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Panger, Gary
;   APPLICANT: Vedvick, Tom
;   APPLICANT: Carter, Darrick
;   APPLICANT: Retter, Marc
;   APPLICANT: Mannion, Jane
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;   TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C9
;   CURRENT APPLICATION NUMBER: US/09/614,124B
;   CURRENT FILING DATE: 2001-07-11
;   NUMBER OF SEQ ID NOS: 1668
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 322
;   LENGTH: 1398
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-614-124B-322

Query Match   6.4%;   Score 141.2; DB 4; Length 1398;
Best Local Similarity 62.7%;   Pred. No. 3.8e-21;
Matches 237; Conservative   0; Mismatches 138; Indels   3; Gaps   1;

Qy   381 AGTCTCCCTGAAGAGGCAGTGAATGGGGTGAATCATTTGACAAACTGCTTTCCCATAG 440
              ||| ||| |||||||| |||| |||||| |||| || ||| | ||||
Db   278 ACTTTCTGCTGCTGAGTAATGCAATGGTCTCAATCTCTGGAAAAACTTCTTGCCAACCA 337

Qy   441 AGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAAGAAATATTGA 500
              |||| ||| ||| ||| ||| |||| ||| || ||| |||||||||||||
Db   338 AACTGGTCAAAATGTCTTTGGAAGTTTCCTAAAGTCTGAATTCAGTGAGGAGAATATTGA 397

Qy   501 ATTTTGGATAGCCTGTGAAGATTCAAGAAAAGCAAGGACCTCAACAAATTCACCTTAA 560
              |||||| || |||| ||| |||| ||||||| ||| |||||| ||||
Db   398 GTTCTGGCTGGCTGTGTGAAGACTATAAGAAAACAGAG--TCTGATCTTTTGCCCTGTAA 454

Qy   561 AGCAAAAGCAATATGATGAGAAATTTATACAGACTGATGCCCCAAAAGAGGTTAACCTTGA 620
           |||||| ||||| ||| ||| | ||| |||| || ||||| || || |||||
Db   455 AGCAGAAGAGATATATAAAGCATTTGTGCATTCAGATGCTGCTAAACAAATCAATATTGA 514

Qy   621 TTTTCACACAAAGAAGTCATTACAAACAGCATCCACCTCACCTTTGA 680
           |||| ||| |||| |||| || |||||| ||| ||||
Db   515 CTTCCGCACTCGAGAATCTCAGCCAAGAAGATTAAAGCACCACCCCCAGTGTTTTGA 574

Qy   681 TGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTATACACGTTTTCTGAA 740
           ||| |||||| || ||| || ||| ||||| ||| ||| || ||||
Db   575 TGAAGCACAAAAGTCATATATACTCTTATGGAAAAGGACTCTTATCCCAGGTTCCTCAA 634

Qy   741 ATCTGACATCTATTTAGA 758
           ||| || ||||| ||| ||||
Db   635 ATCAGATATTTACTTAAA 652

Search completed: August 20, 2004, 12:36:24
Job time : 178.734 secs

```
OM nucleic - nucleic search, using sw model

            GenCore version 5.1.6
            Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:     August 20, 2004, 09:59:05 ; Search time 1035.39 Seconds
                                         (without alignments)
                             10518.000 Million cell updates/sec

Title:         US-09-894-749-1
Perfect score: 2217
Sequence:      1 gaattcggcttccatcctaa........aaaaaaaaagggcggccgc 2217

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:      3228839 seqs, 2456066551 residues

Total number of hits satisfying chosen parameters:    6457678

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Published Applications NA:*
        1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
        2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
        3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
        4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
        5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
        6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
        7:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
        8:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
        9:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
       10:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
       11:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
       12:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
       13:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
       14:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
       15:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
       16:  /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
       17:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
       18:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
       19:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	2217	100.0	2217	9	US-09-894-749-1	Sequence 1, Appli
2	2041.6	92.1	2144	13	US-10-258-371B-19	Sequence 19, Appl
3	1816	81.9	1884	17	US-10-115-635-133	Sequence 133, App
4	1753.2	79.1	1840	13	US-10-258-371B-18	Sequence 18, Appl
5	1448.2	65.3	1486	13	US-10-258-371B-6	Sequence 6, Appli
6	782	35.3	848	9	US-09-867-550-847	Sequence 847, App
7	596	26.9	1164	9	US-09-894-749-3	Sequence 3, Appli
8	388	17.5	736	9	US-09-833-381-1010	Sequence 1010, Ap
9	227.2	10.2	241	13	US-10-258-371B-11	Sequence 11, Appl
10	155.6	7.0	1345	13	US-10-342-887-839	Sequence 839, App
11	155.6	7.0	1345	13	US-10-172-118-839	Sequence 839, App
12	155.6	7.0	1345	17	US-10-305-720-1278	Sequence 1278, Ap
13	155.6	7.0	1345	17	US-10-641-643-1243	Sequence 1243, Ap
14	155.6	7.0	1364	9	US-09-925-300-567	Sequence 567, App
15	155.6	7.0	1381	10	US-09-971-429B-32	Sequence 32, Appl
16	155.6	7.0	1382	10	US-09-971-392-13	Sequence 13, Appl
17	153.2	6.9	744	9	US-09-206-639-2	Sequence 2, Appli
18	153.2	6.9	1923	15	US-10-411-224-55	Sequence 55, Appl
19	153.2	6.9	1923	16	US-10-047-021-55	Sequence 55, Appl
20	153.2	6.9	2076	13	US-10-240-425-345	Sequence 345, App
21	153.2	6.9	2190	17	US-10-641-643-795	Sequence 795, App
22	153.2	6.9	5809	15	US-10-133-013-103	Sequence 103, App
23	152.4	6.9	1629	12	US-10-152-319A-1995	Sequence 1995, Ap
24	152.4	6.9	1629	16	US-10-191-803-244	Sequence 244, App
25	145.2	6.5	543	15	US-10-398-953-2	Sequence 2, Appli
26	145.2	6.5	817	15	US-10-398-953-6	Sequence 6, Appli
27	144.8	6.5	1283	16	US-09-814-353-19578	Sequence 19578, A
28	143.8	6.5	800	16	US-10-305-720-1091	Sequence 1091, Ap
29	143.8	6.5	1531	15	US-10-198-846-9863	Sequence 9863, Ap
30	143.8	6.5	1869	15	US-10-175-523-144	Sequence 144, App
31	143.8	6.5	2752	15	US-10-198-846-13843	Sequence 13843, A
32	143.8	6.5	2934	10	US-09-939-209A-1	Sequence 1, Appli
33	143.8	6.5	3165	14	US-10-044-090-113	Sequence 113, App
34	143.8	6.5	3167	10	US-09-918-624B-6	Sequence 6, Appli
35	141.2	6.4	1329	15	US-10-247-671-107	Sequence 107, App
36	141.2	6.4	1380	9	US-09-925-301-450	Sequence 450, App
37	141.2	6.4	1398	9	US-09-736-457-322	Sequence 322, App
38	141.2	6.4	1398	9	US-09-902-941-322	Sequence 322, App
39	141.2	6.4	1398	9	US-09-849-626-322	Sequence 322, App
40	141.2	6.4	1398	10	US-09-476-300-322	Sequence 322, App
41	141.2	6.4	1398	10	US-09-960-706-870	Sequence 870, App
42	141.2	6.4	1398	10	US-09-873-319-567	Sequence 567, App
43	141.2	6.4	1398	13	US-10-283-017-322	Sequence 322, App
44	141.2	6.4	1398	15	US-10-017-754-322	Sequence 322, App
45	141.2	6.4	1398	15	US-10-113-872-322	Sequence 322, App

ALIGNMENTS

```
RESULT 1
US-09-894-749-1
; Sequence 1, Application US/09894749
; Patent No. US20020081683A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; APPLICANT: Yowe, David
; TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
; FILE REFERENCE: 5800-19, 035800/174680
; CURRENT APPLICATION NUMBER: US/09/894,749
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/244,314
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (160)..(867)
US-09-894-749-1

Query Match             100.0%;  Score 2217;  DB 9;  Length 2217;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 2217;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 GAATTCGGCTTCCATCCTAATAACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAATTCGGCTTCCATCCTAATAACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT 60

Qy     61 ATAACTTTTTATTCTACTATGTATATCGTATGGAAGTAGTATTAATAAATGAACTAGGAAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATAACTTTTTATTCTACTATGTATATCGTATGGAAGTAGTATTAATAAATGAACTAGGAAG 120

Qy    121 GATGTAATAAATTAGACATCTTCATTTTAGAGAAGAGATGGAAACAACATTGCTTTTC 180
```

RESULT 2
US-10-258-371B-19
; Sequence 19, Application US/10258371B
; Publication No. US20040067903A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS-GAGNON, Alison

```
; APPLICANT: MURRAY, David L
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING,
; TITLE OF INVENTION: RGS18, AND USES THEREOF
; FILE REFERENCE: A3656 US PCT
; CURRENT APPLICATION NUMBER: US/10/258,371B
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: GB001883.334
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US60/200,786
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 2144
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-258-371B-19
```

```
Query Match    92.1%;  Score 2041.6;  DB 13;  Length 2144;
Best Local Similarity  99.5%;  Pred. No. 0;
Matches 2058;  Conservative  0;  Mismatches  9;  Indels  1;  Gaps  1;
```

```
Qy   61 ATAACTTTTTATTCTACTATGTATATGTAGGAATAGTATTAATAAATGAACTAGGGAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64 ATAACTTTTTATTCTACTATGTATATGTAGGAATAGTATTAATAAATGAACTAGGGAAG 123

Qy  121 GATGTAATAAATTAGACATCTCTTCATTTGAGAGAGGATGGAAACAACATTGCTTTTC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  124 GATGTAATAAATTAGACATCTCTTCATTTGAGAGAGGATGGAAACAACATTGCTTTTC 183

Qy  181 TTTTTCAAATAAATATGTGAATCAAAAGAAAAAACTTTTTCAAGTAATAACATGGT 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  184 TTTTTCAAATAAATATGTGAATCAAAAGAAAAAACTTTTTCAAGTAATAACATGGT 243

Qy  241 TCAGGAAAAGAAGAAACAGCAAAGAAGCCAAAATCAGAGCTAAGGAAAAAAGAAATAGA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  244 TCAGGAAAAGAAGAAACAGCAAAGAAGCCAAAATCAGAGCTAAGGAAAAAAGAAATAGA 303

Qy  301 CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  304 CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGG 363

Qy  361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  364 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT 423

Qy  421 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTACCAGATTTCTTAAAACTGAA 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  424 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTACCAGATTTCTTAAAACTGAA 483

Qy  481 TTCAGTGAAGAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  484 TTCAGTGAAGAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA 543

Qy  541 CCTCAACAAATTCACCTAAAGCAAAAGCAATATATGAGAAATTTATACAGACTGATGCC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  544 CCTCAACAAATTCACCTAAAGCAAAAGCAATATATGAGAAATTTATACAGACTGATGCC 603

Qy  601 CCAAAAGAGGTTAACCTTGATTTTCACACAGAAGAGTCATTACAAACAGCATCACTCAA 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  604 CCAAAAGAGGTTAACCTTGATTTTCACACAGAAGAGTCATTACAAACAGCATCACTCAA 663

Qy  661 CCTACCTCCACAGTTTGATGCTGCACAAGCTCGACATCTATTGTGCACTTGATGGAAGACCTCAGAACAAGAC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  664 CCTACCTCCACAGTTTGATGCTGCACAAGCTCGACATCTATTGTGCACTTGATGGAAGACCTCAGAACAAGAC 723

Qy  721 AGTTATACACGTTTTCTGAAATCTGACATCTATTTAGACTTGAATGGAAGGAAGACCTCAG 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  724 AGTTATACACGTTTTCTGAAATCTGACATCTATTTAGACTTGAATGGAAGGAAGACCTCAG 783

Qy  781 AGACCAACAAATCTTAGGAGACGATCACGCTCATTACCTGCAATGAATTCCAAGATGTA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  784 AGACCAACAAATCTTAGGAGACGATCACGCTCATTACCTGCAATGAATTCCAAGAGTGTA 843
```

```
Qy  841 CAATCAGATGTTGCCATTGGTTATAAGAAATTATTTTAAAAATGATTTGCTCATTTTATGACAAAC 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  844 CAATCAGATGTTGCCATTGGTTATAAAGAAATTGATTTGCTCATTTTATGACAAAC 903

Qy  901 TTATACATCTGCTTCTAACATATCGCATGTTTATGTTATGTTAAGATTGGTCCCATCCTTTAAA 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  904 TTATACATCTGCTTCTAACATATCGCATGTTTATGTTATGTTAAGATTGGTCCCATCCTTTAAA 963

Qy  961 CTGAAAATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACTTCTGCTAACAAA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  964 CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACTTCTGCTAACAA 1023

Qy 1021 ATACATACAGTATCTGCCAGTATATTCGTAAAACCTTCTATTTGATGTCATTCCATTA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1024 ATACATACAGTATCTGCCAGTATATTCGTAAAACCTTCTATTTGATGTCATTCCATTA 1083

Qy 1081 TAATCAGAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAAGTAATAAATGTTT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1084 TAATCAGAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAAGTAATAAATGTTT 1143

Qy 1141 TATAAGATTGTGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTGTCAAAGTGTCAAACA 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1144 TATAAGATTGTGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTGTCAAAGTGTCAAACA 1203

Qy 1201 AAAGTCAGTTGGGATTTTTTACCAAAGCAGCATAATATGTTATATAAACATAATAAT 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1204 AAAGTCAGTTGGGATTTTTTACCAAAGCAGCATAATATGTTATATAAACATAATAAT 1263

Qy 1261 ACTTCAGATATCCAAATGTTCAGATAGCATTTTTCATAATGAATGTTCTCTTTTTTTGGT 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1264 ACTTCAGATATCCAAATGTTCAGATAGCATTTTTCATAATGAATGTTCTCTTTTTTTGGT 1323

Qy 1321 AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGTGAAGAACATTTATTATTGGGTT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1324 AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGTGAAGAACATTTATTATTGGGTT 1383

Qy 1381 ACTACTACCCTGTCCCAGAATGAGTAATATCACCCTCAGTTGTATAAGCCAGCAACAGGAA 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1384 ACTACTACCCTGTCCCAGAATGAGTAATATCACCCTCAGTTGTATAAGCCAGCAACAGGAA 1443

Qy 1441 CTTTTGTTGAAGACACATTCATCTCTACAGAACTTCAGATTAAATCTAGATTAATG 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1444 CTTTTGTTGAAGACACATTCATCTCTACAGAACTTCAGATTAAATCTAGATTAATG 1503

Qy 1501 ACTGAGAATAAGATCCATCATTTGAACTCATCCCTAAGTGAACGTACCCAGTTAT 1560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1504 ACTGAGAATAAGATCCATCATTTGAACTCATCCCTAAGTGAACGTACCCAGTTAT 1563

Qy 1561 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTTGCATATCTCCAGGTAGGA 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1564 ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTTGCATATCTCCAGGTAGGA 1623

Qy 1621 ACTAAGTAGACTACCTTATCACCGGGCTAAGAGAAACTGCTACTAAACTATTAGGCCATCA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1624 ACTAAGTAGACTACCTTATCACCGGGCTAAGAGAAACTGCTACTAAACTATTAGGCCATCA 1683

Qy 1681 ATGGCTTGAATAAAAACCAGAGAAGGTTTTTCCCAGGACGGTCCATGTTTGGCCCTTTAG 1740
         |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db 1684 ATGGGCTGGAATAAAAACCAGAGAA-GTTTTTCCCAGGACGGTCCATGTTTGGCCCTTTAG 1742

Qy 1741 AATTGGGGTAGAAATCAGAAATGAGATGAGGGGAAGCAAGGAGTCTAAGGCCTAGC 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1743 AATTGGGGTAGAAATCAGAAATGAGATGAGGGGAAGCAAGGAGTCTAAGGCCTAGC 1802

Qy 1801 GATTTGGGCATCTGCCACATTGGTTCATTCATTGATTATTATTC 1860
         ||||||||||||||||||||||||||||||||||||||||||||
Db 1803 GATTTGGGCATCTGCCACATTGGTTCATTCATTGATTATTATTC 1862

Qy 1861 TTGTTAAGCAAATCTCCTTAAGTAATTATTATTCAAATAAGATTATTATTACTCATACACTCTAT 1920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1863 TTGTTAAGCAAATCTCCTTAAGTAATTATTATTCAAATAAGATTATTATTACTCATACATCTAT 1922
```

Qy   1921 ATGTCACTGTTTTAAAGAGATATTTAATTTTTAATGTGTGTTACACGGTCGTGTAAATATT 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1923 ATGTCACTGTTTTAAAGAGATATTTAATTTTTAATGTGTGTTACACGGTCGTGTAAATATT 1982

Qy   1981 TGTATTTAAAAATGCCATGCATTAGGCTTGGAAAATTAATGTAGTTGAAATGTAAAAT 2040
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1983 TGTATTTAAAAATGCCATGCATTAGGCTTGGAAAATTAATGTAGTTGAAATGTAAAAT 2042

Qy   2041 GTGAAAACTTAGATCATCATTTGTAGTAATAAAATATTTTTAACTTCATTCATACAGTTAAGT 2100
           |||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   2043 GTGAAAACTTAGATCATCATTTGTAGTAATAAAATATTTTTAACTTCATTCATACAATTAAGT 2102

Qy   2101 TTATCTGACAATAAAAGCTCTGACTGAA 2128
           ||||||||||||||||||||||||||||
Db   2103 TTATCTGACAATAAAAGCTCTGACTGAA 2130

Db   298 CTTAAGTCTTCTTGTGTGCAGAAACAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGAGTCTGGG 357

Db   361 CACTTGGCCAAGAAACAACAGAGTCTCCCTGAAGAGGCAGTGAAATGGGTGAATCATTT 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   358 CACTTGGCCAAGAAACAACAGAGTCTCCCTGAAGAGGCAGTGAAATGGGTGAATCATTT 417

Qy   421 GACAAACTGCTTCCCATAGAGATGGACTAGAGCCTTTACCAGATTTCTAAAACTGAA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   418 GACAAACTGCTTCCCATAGAGATGGACTAGAGCCTTTACCAGATTTCTAAAACTGAA 477

Qy   481 TTCAGTGAAGAAATATTGAATTTGGATAGCCTGTGAGATTCAAGAAAAGCAAGGGA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   478 TTCAGTGAAGAAATATTGAATTTGGATAGCCTGTGAGATTCAAGAAAAGCAAGGGA 537

Qy   541 CCTCAACAAATTCACCTTAAAGCAAAGCAATATATGAGCAATATATGAGACTGATGCC 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   538 CCTCAACAAATTCACCTTAAAGCAAAGCAATATATGAGCAATATATGAGACTGATGCC 597

Qy   601 CCAAAAGAGGTAACCTTGATTTCCACCAAAAGAAGTCATTACAACACAGCATCACTCAA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   598 CCAAAAGAGGTAACCTTGATTTCCACCAAAAGAAGTCATTACAACACAGCATCACTCAA 657

Qy   661 CCTACCCTCCACAGTTTGATGCTGCACAAGCACAGAGTGTATCAGCTCATGGACAACAAGAC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   658 CCTACCCTCCACAGTTTGATGCTGCACAAGCACAGAGTGTATCAGCTCATGGACAACAAGAC 717

Qy   721 AGTTATACACGTTTCTGAAATCTGACATCATTAGACTGATGGAGGAAGACCTCAG 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   718 AGTTATACACGTTTCTGAAATCTGACATCATTAGACTGATGGAGGAAGACCTCAG 777

Qy   781 AGACCAACAAAATCTTAGGAGACGATCACGCTCATTACCTGCAATGAATTCCAAGATGTA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   778 AGACCAACAAAATCTTAGGAGACGATCACGCTCATTACCTGCAATGAATTCCAAGATGTA 837

Qy   841 CAATCAGATGTTGCCATTGGTATAAAGAGAAAATGAATTTGCTCATTTTTTATGACAAAC 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   838 CAATCAGATGTTGCCATTGGTATAAAGAGAAAATGAATTTGCTCATTTTTTATGACAAAC 897

Qy   901 TTATACATCTGCTTCTAACATCGCATGTTTATGTTAAGATTGTTGTCCCATCCTTTAAA 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   898 TTATACATCTGCTTCTAACATCGCATGTTTATGTTAAGATTGTTGTCCCATCCTTTAAA 957

Qy   961 CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAACAAAACTTTCTGCTAACAAA 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   958 CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAACAAAACTTTCTGCTAACAAA 1017

Qy   1021 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCATTTGATGTCATTCCATTTA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1018 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCATTTGATGTCATTCCATTTA 1077

Qy   1081 TAATCAGAAAAAAACTTATTTCTTAATCAATCAAAGTAAGTTCAAAGTTCAAACA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1078 TAATCAGAAAAAAAAACTTATTTCTTAATCAATCAAAGTAAGTTCAAAGTTCAAACA 1137

Qy   1141 TATAAGATTGTAGAGTTAAGTAAAAGTTAAGCTTTTGCAAAGTTGTCAAAGTTCAAACA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1138 TATAAGATTGTAGAGTTAAGTAAAAGTTAAGCTTTTGCAAAGTTGTCAAAGTTCAAACA 1197

Qy   1201 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAATATGTGTTATAAACATAAT 1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1198 AAAGTCTAGTTGGGATTTTTACCAAAGCAGCATAATATGTGTTATAAACATAAT 1257

Qy   1261 ACTTCAGGATATCCAAATGTTCAGTAATTCAGTAATTAATACATCGT 1320
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1258 ACTTCAGGATATCCAAATGTTCAGTAATTCAGTAATTAATACATGT 1317

Qy   1321 AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAAGAACATTTATTGGGTT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1318 AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAAGAACATTTATTGGGTT 1377

Qy   1381 ACTACTAACCCTGTCCCAAGAATAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAA 1440

RESULT 3
US-10-115-635-133
; Sequence 133, Application US/10115635
; Publication No. US20040137434A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 797CON
; CURRENT APPLICATION NUMBER: US/10/115,635
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 09/714,936
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 362
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 133
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (157)..(864)
US-10-115-635-133

Query Match   81.9%; Score 1816; DB 17; Length 1884;
Best Local Similarity   99.9%; Pred. No. 0; Indels 1; Gaps 1;
Matches 1827; Conservative 0; Mismatches 0; Conservative 0;

Qy   61 ATAACTTTTTATTCTACTATGTATATGTATGGAATAGTATTAATAAATGAACTAGGGAAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   58 ATAACTTTTTATTCTACTATGTATATGTATGGAATAGTATTAATAAATGAACTAGGGAAG 117

Qy   121 GATGTAATAAATTAGACATCTCTTCATTTGAGAGAAGATGGAAACAACATTGCTTTTC 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   118 GATGTAATAAATTAGACATCTCTTCATTTGAGAGAAGATGGAAACAACATTGCTTTTC 177

Qy   181 TTTTTCAAATAAATAATGTGAATCAAAGAAGAAAAACTTTTTCAAGTTAATACATGT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   178 TTTTTCAAATAAATAATGTGAATCAAAGAAGAAAAACTTTTTCAAGTTAATACATGT 237

Qy   241 TCAGGAAAGAAGAAGAACAAGCAAAGAAGCCAAAATCAGAGCTAGGAAAAAAGAAATAGA 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   238 TCAGGAAAGAAGAAGAACAAGCAAAGAAGCCAAAATCAGAGCTAGGAAAAAAGAAATAGA 297

Qy   301 CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGAGTCTGGG 360

Db   1378 ACTACTAACCGTGTCCCAAGAATAAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAA 1437
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184 TTTTCTCAAATAAATATGTGTGAATCAAAGAAAGAAAAACTTTTTTCAAGTAATACATGGT 243

Qy   1441 CTTTTGTGAAGACACATTCATCTCTACAGAACTTCAGATTAATAATATTAGATTAATG 1500
          |||||||||||||||||||||||||||||||||||||||||||||
Qy   241 TCAGGAAAGAGAAGAGAAGAAACAAGACAAATCAGAGCTAAGGAAAAATCAGAGAAATAGA 300

Db   1438 CTTTTGTGAAGACACATTCATCTCTACAGAACTTCAGATTAAATAATATTAGATTAATG 1497
          |||||||||||||||||||||||||||||||||||||||||||||
Db   244 TCAGGAAAGAGAAGAGAAGAAACAAGACAAATCAGAGCTAAGGAAAAAAGAAAATAGA 303

Qy   1501 ACTGAGAATAAGATCCACATTGAACTCATTCATTCCTCTAAGTGAACATGGACGTACCAGTTAT 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   301 CTAAGTCTTCGTGTGCAGAAACAAGAGTCTCCCCTGAGAGACAGGCAGTGAAATGGGGTCGG 360

Db   1498 ACTGAGAATAAGATCCACATTGAACTCATTCATTCCTCTAAGTGAACATGGACGTACCAGTTAT 1557
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   304 CTAAGTCTTCGTGTGCAGAAACAAGAGTCTCCCCTGAGAGACAGGCAGTGAAATGGGGTGGG 363

Qy   1561 ACAAAGTACTTCGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGGTAGGGA 1620
          ||||||||||||||||||||||||||||||||||||||||||||
Qy   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAGAGGGTGAATCATTT 420

Db   1558 ACAAAGTACTTCGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGGTAGGGA 1617
          ||||||||||||||||||||||||||||||||||||||||||||
Db   364 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAGAGGGTGAATCATTT 423

Qy   1621 ACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTGCTACTAAACTATTAGGCCATCA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||
Qy   421 GACAAACTGCTTTCCCATAGAATGGACTAGAGGCTTTACCAGATTTCTTAAAACTGAA 480

Db   1618 ACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTGCTACTAAACTATTAGGCCATCA 1677
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   424 GACAAACTGCTTTCCCATAGAATGGACTAGAGGCTTTACCAGATTTCTTAAAACTGAA 483

Qy   1681 ATGGCTTGAATAAAAACCAGAGAGAGGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG 1740
          ||||||||||||||||||||||||||||||||||||||||||||
Qy   481 TTCAGTGAAGAAATATTGAATTTGGATAGGCTGTGAGAGATTCAAGAAAAGCAAGGGA 540

Db   1678 ATGGCTTGAATAAAAACCAGAGAGAGGTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG 1737
          ||||||||||||||||||||||||||||||||||||||||||||
Db   484 TTCAGTGAAGAAATATTGAATTTGGATAGGCTGTGAGAGATTCAAGAAAAGCAAGGGA 543

Qy   1741 AATTGGGGTAGAAATCAGAAATGAGAATGAGGGGAAGAAGCAAGGAGTCTAAGGCCCTAGC 1800
          ||||||||||||||||||||||||||||||||||||||||||||
Qy   541 CCTCAACAAATTCACCTTAAGCAAAGCAATAATGAGCAATATTATACCACAGCATGATGCC 600

Db   1738 AATTGGGGTAGAAATCAGAAATGAGAATGAGGGGAAGAAGCAAGGAGTCTAAGGCCCTAGC 1797
          ||||||||||||||||||||||||||||||||||||||||||||
Db   544 CCTCAACAAATTCACCTTAAGCAAAGCAATAATGAGCAATATTATACCACAGCATGATGCC 603

Qy   1801 GATTTGGGCCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCCATTGATTATATTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||
Qy   601 CAAAAGAGGTTAACCTTGATTTGATCGTGACAAAGCAGAGTCATCAGAACAACAAGAC 660

Db   1798 GATTTGGGCCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCCATTGATTATATTC 1857
          ||||||||||||||||||||||||||||||||||||||||||||
Db   604 CAAAAGAGGTTAACCTTGATTTGATCGTGACAAAGCAGAGTCATCAGAACAACAAGAC 663

Qy   1861 TTGTTAAGCAAATCTCCTTAAGTAATTA 1888
          ||||||||||||||||||||||
Qy   661 CCTACCCTCCACAGTTTGATGCTGACAAAGCAGTGTATCAGCTCATGGAAGGAAGACCTCAG 720

Db   1858 TTGTTAAGC-AATCTCCTTAAGTAATTA 1884
          ||||||||||||||||||||||
Db   664 CCTACCCTCCACAGTTTGATGCTGACAAAGCAGTGTATCAGCTCATGGAAGGAAGACCTCAG 723

Qy   721 AGTTATACACGTTTTCTGAAATCTGACATCTATTAGACTTGACTTGATGGAAGGAAGACCTCAG 780
          ||||||||||||||||||||||||||||||||||||||||||||
Qy   724 AGTTATACACGTTTTCTGAAATCTGACATCTATTAGACTTGACTTGATGGAAGGAAGACCTCAG 783

Qy   781 AGACCACAAAATCTTAGGAGACGATCACGCTCATTCACCTGCAATGAATTCCAAGATGTA 840
          ||||||||||||||||||||||||||||||||||||||||||||
Db   784 AGACCACAAAATCTTAGGAGACGATCACGCTCATTCACCTGCAATGAATTCCAAGATGTA 843

Qy   841 CAATCAGATGTTGCCATTTGGTTATAAAGAAAATTGATTTTGCTCATTTTTATGACAAAC 900
          ||||||||||||||||||||||||||||||||||||||||||||
Db   844 CAATCAGATGTTGCCATTTGGTTATAAAGAAAATTGATTTTGCTCATTTTTATGACAAAC 903

Qy   901 TTATACATCTGCTTCAACATATCGCCAGTATATCGTAAGTTAATGTTAAGATTTGGTCCCATCCTTTAAA 960
          ||||||||||||||||||||||||||||||||||||||||||||
Db   904 TTATACATCTGCTTCAACATATCGCCAGTATATCGTAAGTTAATGTTAAGATTTGGTCCCATCCTTTAAA 963

Qy   961 CTGAAATGTCATGTGAAAATTATTTTAAAAATGTAAAAATAACAAAACTTCTGTCATTCATTA 1020
          ||||||||||||||||||||||||||||||||||||||||||||
Db   964 CTGAAATGTCATGTGAAAATTATTTTAAAAATGTAAAAATAACAAAACTTCTGTCATTCATTA 1023

Qy   1021 ATACATACAGTATCTGCCAGTATATTCTTAATCAAAAAGTACAAAAAAGTAATAATGTTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1024 ATACATACAGTATCTGCCAGTATATTCTTAATCAAAAAGTACAAAAAAGTAATAATGTTT 1083

Qy   1081 TAATCAGAAAAAAACTTATTTCTTAATCAAAAAGTACAAAAAAGTAATAATGTTT 1140
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1084 TAATCAGAAAAAAACTTATTTCTTAATCAAAAAGTACAAAAAAGTAATAATGTTT 1143

Qy   1141 TATAAGATTGTAGAGTTAAGTAAAAGTTAAGCTTTGCAAAAGTTGTCAAAAGTTCAAACA 1200
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1144 TATAAGATTGTAGAGTTAAGTAAAAGTTAAGCTTTGCAAAAGTTGTCAAAAGTTCAAACA 1203

Db   61 ATAACTTTTTATTCTACTATGTATATGTATATGTATATAATAGTATTAATAAATGAACTAGGGAAG 120
          ||||||||||||||||||||||||||||||||||||||||||||
Db   64 ATAACTTTTTATTCTACTATGTATATGTATATGTATATAATAGTATTAATAAATGAACTAGGGAAG 123

Qy   1201 AAAGTCTAGTTGGGGATTTTTACCAAGACAGCAGATGGAAACAACATTGCTTTTC 1260
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1204 AAAGTCTAGTTGGGGATTTTTTACCAAGACAGCAGCAGATGGAAACAACATTGCTTTTC 1263

Qy   121 GATGTAATAAATTAGACATCTCTTCATTTAGAGACGAAGATGGAAACAACATTGCTTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||
Db   124 GATGTAATAAATTAGACATCTCTTCATTTAGAGACGAAGATGGAAACAACATTGCTTTTC 183

Qy   1261 ACTCAGATATCCAAATGTTCAGATAGCATTTTCATAAGAATGATGTGTCTCTTTTTTTTGGT 1320
          ||||||||||||||||||||||||||||||||||||||||||||
Qy   1264 ACTCAGATATCCAAATGTTCAGATAGCATTTTCATAAGAATGATGTGTCTCTTTTTTTTGGT 1323

Qy   181 TTTTCTCAAATAAATATGTGTGAATCAAAGAAAGAAAAACTTTTTTCAAGTAATACATGGT 240
          ||||||||||||||||||||||||||||||||||||||||||||

RESULT 4
US-10-258-371B-18
; Sequence 18, Application US/10258371B
; Publication No. US20040067903A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS-GAGNON, Alison
; APPLICANT: MURRAY, David L
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING,
; TITLE OF INVENTION: RGS18, AND USES THEREOF
; FILE REFERENCE: A3656 US PCT
; CURRENT APPLICATION NUMBER: US/10/258,371B
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: GB001883.334
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US60/200,786
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-258-371B-18

Query Match   79.1%; Score 1753.2; DB 13; Length 1840;
Best Local Similarity 99.5%;
Matches 1769; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy   776  CTCAGAGACCAACAAATCTTAGGAGGACGATCACGCTCATTCACCTGCAATGAATTCCAAG  835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  CTCAGAGACCAACAAATCTTAGGAGGACGATCACGCTCATTCACCTGCAATGAATTCCAAG  180

Qy   836  ATGTACAATCAGATGTTGCCATTTGCTTATAAAGAAAATGATTTGCTCATTTTTATGA  895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  ATGTACAATCAGATGTTGCCATTTGCTTATAAAGAAAATGATTTGCTCATTTTTATGA  240

Qy   896  CAAACTTATACATCTGCTTCTCAACATATCGCATGTTTATGTTAAGATTTGGTCCCATCCT  955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  CAAACTTATACATCTGCTTCTCAACATATCGCATGTTTATGTTAAGATTTGGTCCCATCCT  300

Qy   956  TTAAACTGAAATATGTCATGTGAAATTATTTAAAAATGTAAAAACAAAACTTTCTGCTA  1015
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  TTAAACTGAAATATGTCATGTGAAATTATTTAAAAATGTAAAAACAAAACTTTCTGCTA  360

Qy  1016  ACAAAATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCATTTGATGTCATTCC  1075
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  ACAAAATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCATTTGATGTCATTCC  420

Qy  1076  ATTTATAATCAGAAAAAAAACTATTTCTTAATCAAACTATTAGGCAGTACAAAAAAGTAATAA  1135
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  ATTTATAATCAGAAAAAAAACTATTTCTTAATCAAACTATTAGGCAGTACAAAAAAGTAATAA  480

Qy  1136  TGTTTTATAAGATTGTAGAGTTAAGTAAGAGTTAAGTAAAAGTTGTCAAAAGTTC  1195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TGTTTTATAAGATTGTAGAGTTAAGTAAGAGTTAAGTAAAAGTTGTCAAAAGTTC  540

Qy  1196  AACAAAAGTCTAGTGGATTTTTACCAAAGCAGCATATATGTTATATAAACATA  1255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  AACAAAAGTCTAGTGGATTTTTACCAAAGCAGCATATATGTTATATAAACATA  600

Qy  1256  ATAATACTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTTT  1315
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  ATAATACTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCTCTTTTTT  660

Qy  1316  TTGGTAATAGTGTAGAAGTGATCTCGTTCTTACAATGGGAGGATGAAGAACATTATTATT  1375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  TTGGTAATAGTGTAGAAGTGATCTCGTTCTTACAATGGGAGGATGAAGAAGACATTATTATT  720

Qy  1376  GGGTTACTACTAACCCTGTCCAAGAATAGTAATATCACCTCTAGTTATAAGCCAGCAAC  1435
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  GGGTTACTACTAACCCTGTCCAAGAATAGTAATATCACCTCTAGTTATAAGCCAGCAAC  780

Qy  1436  AGGAACTTTGTGAAGACACATTCATCTCACGAACTTCAGATTAAATAATCTAGAT  1495
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  AGGAACTTTGTGAAGACACATTCATCTCACGAACTTCAGATTAAATAATCTAGAT  840

Qy  1496  TAATGACTGAGAATAGATCCACATTGAACTCATTCCTAAGTGACGTACCCA  1555
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  TAATGACTGAGAATAGATCCACATTGAACTCATTCCTAAGTGACGTACCCA  900

Qy  1556  GTTATACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGGT  1615
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  GTTATACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGGT  960

Qy  1616  AGGGACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACATTAGGC  1675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  AGGGACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACATTAGGC  1020

Qy  1676  CATCAATGGCTTGAATAACCAGAAAAACCAGAA-GTTTTTCCCAGGAGTCTCATGTTTGGCCC  1735
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  1021  CATCAATGGCTGGAATAACCAGAA-GTTTTTCCCAGGAGTCTCATGTTTGGCCC  1079

Qy  1736  TTTAGAATTGGGTAGAAATCAGAAATGCTGCACAAAGCAGAGTGTATCAAGGCTCTAAGGCC  1795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1080  TTTAGAATTGGGTAGAAATCAGAAATGCTGCACAAAGCAGAGTGTATCAAGGCTCTAAGGCC  1139

Qy  1796  CTAGCCATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGATTA  1855
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1140  CTAGCCATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTGTTATCTCATTGATTA  1199
```

```
Qy  1321  AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1324  AATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTT  1383

Qy  1381  ACTACTAACCCTGTCCAAGAATAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAA  1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1384  ACTACTAACCCTGTCCAAGAATAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAA  1443

Qy  1441  CTTTGTGAAGACACATTCATCTCACAGAACTTCAGATTAAATAATCTAGATTAATG  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1444  CTTTGTGAAGACACATTCATCTCACAGAACTTCAGATTAAATAATCTAGATTAATG  1503

Qy  1501  ACTGAGAATAAGATCCACATTCCTAAGTGAACATGACGTACCCAGTTAT  1560
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1504  ACTGAGAATAAGATCCACATTCCTAAGTGAACATGGACGTACCCAGTTAT  1563

Qy  1561  ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGGTAGGGA  1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1564  ACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGGTAGGGA  1623

Qy  1621  ACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGCCCATCA  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1624  ACTAAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTAGGCCCATCA  1683

Qy  1681  ATGGCTTGAATAAAAACCAGAAGGTTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1684  ATGGGCTGGAATAAAAACCGAGAA-GTTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAG  1742

Qy  1741  AATTGGGTAGAAATCAGAAATGAGGATGAGGGGAAGCAAGGAGTCCTAAGGCCCTAGC  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1743  AATTGGGTAGAAATCAGAAATGAGGATGAGGGGAAGCAAGGAGTCCTAAGGCCCTAGC  1802

Qy  1801  GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAG  1838
          ||||||||||||||||||||||||||||||||||||||
Db  1803  GATTTGGGCATCTGCCACATTGGTTCATATTCAGAAAG  1840
```

RESULT 5
US-10-258-371B-6
; Sequence 6, Application US/10258371B
; Publication No. US20040067903A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS-GAGNON, Alison
; APPLICANT: MURRAY, David L
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING,
; TITLE OF INVENTION: RGS18, AND USES THEREOF
; FILE REFERENCE: A3656 US PCT
; CURRENT APPLICATION NUMBER: US/10/258,371B
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: GB001883.334
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US60/200,786
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-258-371B-6

Query Match   65.3%; Score 1448.2; DB 13; Length 1486;
Best Local Similarity   99.4%; Pred. No. 8.1e-274;
Matches 1464; Conservative   0; Mismatches   8; Indels   1; Gaps   1;

```
Qy   656  CTCAACCTACCCTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATCATGGAAC  715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  CTCAACCTACCCTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATCATGGAAC  60

Qy   716  AAGACAGTTATACACGTTTCTGAAATCTGACATCTATTTAGACTTGACTTGATGGAGGAAGAC  775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  AAGACAGTTATACACGTTTCTGAAATCTGACATCTATTTAGACTTGACTTGATGGAGGAAGAC  120
```

```
Qy  1856  TATTCCTTGGTTAAGCAAATCTCCTTAAGTAATATTATTCAAATAAGAGTATACTCATACA  1915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1200  TATTCCTTGGTTAAGCAAATCTCCTTAAGTAATATTATTCAAATAAGAGTATACTCATACA  1259

Qy  1916  TCTATATGTCACTCGTTTTAAAGAGATATTTAATTTTAATGTGTGTTACATGGTCTGTAA   1975
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1260  TCTATATGTCACTCGTTTTAAAGAGATATTTAATTTTAATGTGTGTTACATGGTCTGTAA   1319

Qy  1976  ATATTTGTATTTAAAAAATGCCATGCATTAGCCATTGCATTTGGAAATTTAATGTTAGTTGAAAATGT  2035
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1320  ATATTTGTATTTAAAAAATGCCATGCATTAGCCATTGCATTTGGAAATTTAATGTTAGTTGAAAATGT  1379

Qy  2036  AAAATGTGAAAACTTAGATCATTTGTAGTAATAAATATTTTTAACTTCATTCATCATACAGT  2095
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||| ||||||
Db  1380  AAAATGTGAAAACTTTAGATCATTTGTAGTAATAAATATTTTTAACTTCATTCATCATACAAT  1439

Qy  2096  TAAGTTTATCTGACAATAAAAGCTCTGACTGAA  2128
          |||||||||||||||||||||||||||||||||
Db  1440  TAAGTTTATCTGACAATAAAAGCTCTGACTGAA  1472
```

RESULT 6
US-09-867-550-847
; Sequence 847, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1e1 Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; SOFTWARE: FastSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 2125
; SEQ ID NO 847
; LENGTH: 848
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: wherein any n is one of a or t or g or c
US-09-867-550-847

Query Match   35.3%; Score 782; DB 9; Length 848;
Best Local Similarity 100.0%; Pred. No. 2.6e-143;
Matches   782; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    61  ATAACTTTTTATTCTACTATGTATATGTATGGAATAGTATTAATAATGAACTAGGGAAG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67  ATAACTTTTTATTCTACTATGTATATGTATGGAATAGTATTAATAATGAACTAGGGAAG  126

Qy   121  GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127  GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC  186

Qy   181  TTTTCTCAAATAAATAAATGTGTGAATCAAAAGAAAAAACCTTTTTCAGTGAATAACATGGT  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187  TTTTCTCAAATAAATAAATGTGTGAATCAAAAGAAAAAACCTTTTTCAGTGAATAACATGGT  246

Qy   241  TCAGGAAAAGAAGAAGAACAAGCAAAGAAGCCAAATCAGAGCTAAGGAAAAAGAAATAGA  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247  TCAGGAAAAGAAGAAGAACAAGCAAAGAAGCCAAATCAGAGCTAAGGAAAAAGAAATAGA  306

Qy   301  CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGG  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307  CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGG  366
```

```
Qy   361  CACTTGGCCAAAGAACAAGAGTCTCCCTGAAGAGGCAGTGAAATGGGTGAATCATTT  420
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   367  CACTTGGCCAAAGAAACAAGAGTCTCCCTGAAGAGGCAGTGAAATGGGTGAATCATTT  426

Qy   421  GACAAATGCTTTCCCATAGAGATGGACTAGAGGCTTTACCAGATTTCTAAAACTGAA  480
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   427  GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTACCAGATTTCTAAAACTGAA  486

Qy   481  TTCAGTGAGAAAATATGAATTTGGATAGCCTGTGAGAGATTTCAGAAAAGCAAGGGA  540
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   487  TTCAGTGAGAAAATATTGAATTTGGATAGCCTGTGAGAGATTTCAGAAAAGCAAGGGA  546

Qy   541  CCTCAACAAATCACCTTAAAGCAATATATGAGAAATTTATACAAACAGCATCACTCAA  600
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   547  CCTCAACAAATTCACCTTAAAGCAATATATGAGAAATTTATACAAACAGCATCACTCAA  606

Qy   601  CCAAAAGAGGTAACCTTGATTTTCACACAAAGCAGAGTGTATCAGCTCATGGAACAGAC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607  CCAAAAGAGGTAACCTTGATTTTCACACAAAGCAGAGTGTATCAGCTCATGGAACAGAC  666

Qy   661  CCTACCCTCCACAGTTTGATGCTGCACAAAGCAGAGTGCGACAAACAGAC  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   667  CCTACCCTCCACAGTTTGATGCTGCACAAAGCAGAGTGCGACAAACAGAC  726

Qy   721  AGTTATACACGTTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAGGAGACCTCAG  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   727  AGTTATACACGTTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAGGAGACCTCAG  786

Qy   781  AGACCAACAAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   787  AGACCAACAAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTA  846

Qy   841  CA  842
          ||
Db   847  CA  848
```

RESULT 7
US-09-894-749-3
; Sequence 3, Application US/09894749
; Patent No. US2002008168 3A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; APPLICANT: Yowe, David
; TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
; FILE REFERENCE: 5800-19, 035800/174680
; CURRENT APPLICATION NUMBER: US/09/894,749
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/244,314
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1164
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (134)..(841)
US-09-894-749-3

Query Match   26.9%; Score 596; DB 9; Length 1164;
Best Local Similarity 75.8%; Pred. No. 8.8e-107;
Matches   791; Conservative   0; Mismatches   220; Indels   33; Gaps   3;

```
Qy    69  TTATTCTACTATGTATATGTATGGAATAGTATTAAATGAACTAGGGAAGGATGTAAT  128
          || ||||| |||||| ||||||||||| ||| |||| |||||| ||||| ||||
Db    41  TTTATTCAGGATGTTTTCCTATGAAATAGCATTCATCTGTGGGAGAGAGGACTAAGG  100

Qy   129  AAATTAGACATCTCTTCAT--TTTAGAGAGAAGATGGAAACAACATTGCTTTTCTTCT  186
          ||| ||||||||  || |      || |||| ||| ||||| || |||  |||||||||
Db   101  AAATCTGACATCGTTGGTCACTGGGACAGAATATGGATATGTCACTGGTTTTCTTCTCT  160
```

```
Query Match       17.5%;   Score 388;   DB 9;   Length 736;
Best Local Similarity   100.0%;   Pred. No. 4.2e-66;
Matches 388;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   481 TTCAGTGAAGAAAATATTGAATTTGATAGCCTGTGAAGATTCAAGAAAAGCAAGGGA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4 TTCAGTGAAGAAAATATTGAATTTGATAGCCTGTGAAGATTCAAGAAAAGCAAGGGA 63

Qy   541 CCTCAACAAATTCACCTTAAAGCAAAGCAATATATGAGAAATTTATACAGACTGATGCC 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    64 CCTCAACAAATTCACCTTAAAGCAAAGCAATATATGAGAAATTTATACAGACTGATGCC 123

Qy   601 CCAAAAGAGGTGAACCTTGATTTTCACACAAAAGAAGTCATTACAACAGCATCACTCAA 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   124 CCAAAAGAGGTAACCTTGATTTTCACACAAAAGAAGTCATTACAACAGCATCACTCAA 183

Qy   661 CCTACCCTCCACAGTTTGATGCTGCACAAAGCAGCAGAGTGTATCAGCTCATGGAACAAGAC 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184 CCTACCCTCCACAGTTTGATGCTGCACAAAGCAGCAGAGTGTATCAGCTCATGGAACAAGAC 243

Qy   721 AGTTATACACGTTTCTGAAATCTGACATCTATTAGACTTGATGGAAGGAAGAACCTCAG 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   244 AGTTATACACGTTTCTGAAATCTGACATCTATTAGACTTGATGGAAGGAAGAACCTCAG 303

Qy   781 AGACCAACAAATCTTAGGAGACGATCACGCTCATTACCTGCAATGAATTCCAAGATGTA 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   304 AGACCAACAAATCTTAGGAGACGATCACGCTCATTACCTGCAATGAATTCCAAGATGTA 363

Qy   841 CAATCAGATGTTGCCATTTGTTATAAA 868
         |||||||||||||||||||||||||||
Db   364 CAATCAGATGTTGCCATTTGTTATAAA 391
```

RESULT 9
US-10-258-371B-11

```
Qy   187 CAAATAAATATGTGAATCAAAAGAAAAACTTTTTCAAGTTAATACATGTTCAGGA 246
         |||||||||||||||||||||||||       |||||||||  |||||||  ||||
Db   161 CAATTAAATATGTGAATCAAAAGAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGG 220

Qy   247 AAAGAAGAACAAGCAAGCCAAAATCAGAGCTAAGAGAAAAGAAATAGACTAAGT 306
         |||||||||||||||  ||||||||||||  |||||||||||||||||||||||
Db   221 AAAGAAGAACAAGCAGCCAAAATCAGAGCGAAAATCAGAGCGAAAGAAATAGACTAAGT 280

Qy   307 CTTCTGTGCAGAAACCTGAGTTCATGAAGACACCCGCCTCCGTAGATCTGGGCACTTG 366
         ||||||  ||||||||||||||  |||||||  ||  |||||||||||||||||||||
Db   281 CTTCTCCACAGAGGCCTGACTTCCATGGAGACCAGCAGCTCAAGCCAGTAGATCTGCCCTCTTG 340

Qy   367 GCCAAGAAACAAGAGTCTCCCTGAAGACGCAGTGAAATGGGTGAATCATTTGACAAA 426
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   341 GCCAAGAAACAAGAGTCTCCTGAAGAAGCAGTGAAATGGGTGAATCATTTGACAAA 400

Qy   427 CTGCTTTCCATAGAGATGGACTAGAGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 486
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   401 TTGCTCTCTCATAGAGATGGAGTGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 460

Qy   487 GAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTCAAGAAAAGCAGGACCTCAA 546
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   461 GAGGAGAACATTGAATTTGGGTCGCCTGTGAAGACTTCAAGAAAATGCAAGGAACCTCAA 520

Qy   547 CAAATTCACCTTAAAGCAAAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAA 606
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   521 CAAATCATCCTAAAGCAAAGGCAATCTCATGAGAAATTCATTCAGAATGAGGCTGCCCCCAAA 580

Qy   607 GAGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAAAACAGCATCACTCACCTACC 666
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   581 GAGGTTAACATTTCATCAGAAGTAATTGCTAAAGAAGTGCTAAGAGCATCATCTACCTACC 640

Qy   667 CTCCACAGTTTGATGCTGCACAAAGCAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTAT 726
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   641 CTCCACAGTTTGATGGCACCACGGGCACCTACTTACAGTTACCAGTGTCATTGGGAACATGACAGTTAT 700

Qy   727 ACACGTTTCTGAAATCTGACATCTATTAGACTTGATGGAGGACCA 786
         |||||||||||||||||||||||||||||||||||||||||||||
Db   701 AAACGCTTTTTGAAATCTGAGAACCTACATTTGACAAGCTTACAGGAGACCA 760

Qy   787 ACAAATCTTAGGAGACGATCACGCTCATTACCTGCAATGACATCA 846
         |||||||||||||||||||||||||||||||  |||||||||||||
Db   761 ACAAACCTTAGGAGAACGACGATCACGATCATTACTTACCAACGATGTAAAGTCA 820

Qy   847 GATGTTGCCATTTGGTTATAAAAGAAAATTGATTTTGCTCATTTTTGATGACAAACTTATAC 906
         ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   821 GATGTTGCCATTTGGTTATGAGTAAAAGTCATTTGTCTCTTTCTGATAGTGTATGTGTAT 880

Qy   907 ATC---------------------------------TGCTTCTAACAATATCGCCATGTTATGTGTTAAGA 941
         |||                                 ||||||||||||||  ||  ||||||||||
Db   881 ATCTAAAATATATACTAATACTACTTCTAAAATATAGCTTGTGTTGTATAAATATAGCTTGTGTATAAGAGA 940

Qy   942 TTTGGTCCCATCCTTAAACTGAAATATGTCATGTGAAATATTATTTTAAAAATGTAAAAAC 1001
         ||||||||||||||||||||||||||     |||||||  |||||||||||
Db   941 TTTGGTCCCATCCTTAAACTGAAATATGT------TAAAATACACCATGCAAATAACAATTCAAAATTACAAAATTTAAATGTAAGAACTT 994

Qy  1002 AAAACTTCTGCTAACAAAATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTA 1061
         |||||                    ||||||||||||||||||||||||||
Db   995 TTTTATTATACTAAAAATAATTCATCATCTATCTTCCCAAAATATTTTATGAAAAATCTATC 1054

Qy  1062 TTTGATGTGTCATTCCATTTATAATC 1085
         |||||| | |||||||||||||
Db  1055 TGATATTCTATTCTAATTAAAATTC 1078
```

RESULT 8
US-09-833-381-1010

```
Query Match           10.2%;   Score 227.2;  DB 13;  Length 241;
Best Local Similarity  96.7%;  Pred. No. 8.2e-35;
Matches 232;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy   487  GAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTCAAGAAAAGCAAGGACCTCAA 546
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  GAGGAAAATCTGGAGTTCGGATAGCCTGTGAAGATTCAAGAAAAGCAAGGACCTCAA 60

Qy   547  CAAATTCACCTTAAAGCAAAAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAA 606
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CAAATTCACCTTAAAGCAAAAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAA 120

Qy   607  GAGGTTAACCTTGATTTTCACCAAAGAAGTCATTCATTACAAACAGCATCACTCAACCTACC 666
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GAGGTTAACCTTGATTTTCACCACAAGAAGTCATTCATTACAAACAGCATCACTCAACCTACC 180

Qy   667  CTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTAT 726
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  CTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAAAACGACAGCAGTAT 240
```

RESULT 10
US-10-342-887-839
; Sequence 839, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 839
; LENGTH: 1345
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-839

```
Query Match            7.0%;   Score 155.6;  DB 13;  Length 1345;
Best Local Similarity  63.2%;  Pred. No. 2.1e-20;
Matches 239;  Conservative  0;  Mismatches 139;  Indels  0;  Gaps  0;

Qy   385  TCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTTGACAAACTGCTTTCCCATAGAGAT 444
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   249  TCTCCTGAGGAGCACACAGCTGTGTGGTCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy   445  GGACTAGAGGCTTTTACCAGATTTCTTAAACTGAATTCAGTGAAGAAAATATTGAATTT 504
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   309  GGTCTTGCTGCATTCAGGGCTTTTTTAAAGTCGGAATTCTGTGAAGAAAATATTGAATTC 368

Qy   505  TGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAAAGCA 564
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   369  TGGCTGGCCTGTGAAGACTTCAAAAAACCAATCACCCCAAAAGCTGTCCTCAAAAGCA 428

Qy   565  AAAGCAATATGAGAAATTTATACAGACTGATGCCCCAAAAGAGGTTAACCTTGATTTT 624
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   429  AGGAAAATATATACTGACTTCATAGAAAAGGAAGCTCCAAAGAGAGATAAAACATAGATTTT 488

Qy   625  CACACAAAGAAGTCATTACAAACAGCATCACTCAACCTACCCCTCAACCTACCCCTACCCTCAACC 684
```

RESULT 12
US-10-305-720-1278
; Sequence 1278, Application US/10305720
; Publication No. US2004010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expressio
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1278
; LENGTH: 1345
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. US2004010136A1 g292054
US-10-305-720-1278

Query Match   7.0%;   Score 155.6; DB 16;   Length 1345;
Best Local Similarity 63.2%;   Pred. No. 2.1e-20;
Matches 239; Conservative   0; Mismatches 139; Indels   0; Gaps   0;

```
Qy   385 TCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTTGACAAACTGCTTTCCCATAGAGAT 444
         ||||||||| ||||| |  |||||  |||| ||||||| ||| ||||
Db   249 TCTCCTGAGGAAGCACAGCTGTGGTCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy   445 GGACTAGAGGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAAGAAAATATTGAATTT 504
         || ||  | |||  ||||  | ||||||| |||||||  |||||||||||||||||||
Db   309 GGTCTTGCTGCATTCAGGGCTTTTTAAAGTCCGAATTCTGTGAAGAAAATATTGAATTC 368

Qy   505 TGGATAGCCTGTGAAGATTCAAGAAAAGCAAGGACCTCAACAAATCACCTTAAAGCA 564
         ||| | ||||||||||| ||||| |||| ||| |  ||| |||| ||| || ||||
Db   369 TGGCTGGCCTGTGAAGACTTCAAAAAAAACCAAATCACCCCAAAAAGCTGTCCTCAAAAGCA 428

Qy   565 AAAGCAATATAGAGAAATTATACAGACTGATGCCCCAAAAGAGGTTAACCTTGATTTT 624
         |  |  |||||| | |  ||||| || |||| ||  ||||||| ||||  |||| |||
Db   429 AGGAAAATATATACTGACTTCATAGAAAAGGAAGCTCCAAAGAGATAAACATAGATTTT 488

Qy   625 CACACAAAGAAGTCATTACAAACAGCATCCACTCACTCACTCCACAGTTTTGTGATGCT 684
         ||  | ||| | |  |   ||| |||| ||||  |||  ||||||||||| ||| |||
Db   489 CAAACCAAAACTCTGATTGCCCAGAATATACAAGAAGCTACAAGTGGCTGCTTTACAACT 548

Qy   685 GCACAAAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTATACACGTTTTCTGAAATCT 744
         || ||||| | | ||||| |||| |||| ||| || |   |||| | ||| | ||| ||
Db   549 GCCCAGAAAGGGTATACAGCTTGATGGAGAACAACTCTTATCCTCGTTTCGTTCTTGGAGTCA 608

Qy   745 GACATCTATTTAGACTTG 762
         || ||| || ||||||||
Db   609 GAATTCTTACCAGGACTTG 626
```

RESULT 13
US-10-641-643-1243
; Sequence 1243, Application US/10641643
; Publication No. US2004077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
;   Susan G. Stuart
;   Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
                      GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641,643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1243:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1345 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g292054
; SEQUENCE DESCRIPTION: SEQ ID NO: 1243 :
US-10-641-643-1243

Query Match   7.0%;   Score 155.6; DB 17;   Length 1345;
Best Local Similarity 63.2%;   Pred. No. 2.1e-20;
Matches 239; Conservative   0; Mismatches 139; Indels   0; Gaps   0;

```
Qy   385 TCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTTGACAAACTGCTTTCCCATAGAGAT 444
         ||||||||| ||||| |  |||||  |||| ||||||| ||| ||||
Db   249 TCTCCTGAGGAAGCACAGCTGTGGTCAGAAGCATTTGACGAGCATTTGACGAGCAGCAAATAT 308

Qy   445 GGACTAGAGGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAAGAAAATATTGAATTT 504
         || ||  | |||  ||||  | ||||||| |||||||  |||||||||||||||||||
Db   309 GGTCTTGCTGCATTCAGGGCTTTTTAAAGTCCGAATTCTGTGAAGAAAATATTGAATTC 368

Qy   505 TGGATAGCCTGTGAAGATTCAAGAAAAGCAAGGACCTCAACAAATCACCTTAAAGCA 564
         ||| | ||||||||||| ||||| |||| ||| |  ||| |||| ||| || ||||
Db   369 TGGCTGGCCTGTGAAGACTTCAAAAAAAACCAAATCACCCCAAAAAGCTGTCCTCAAAAGCA 428

Qy   565 AAAGCAATATAGAGAAATTATACAGACTGATGCCCCAAAAGAGGTTAACCTTGATTTT 624
         |  |  |||||| | |  ||||| || |||| ||  ||||||| ||||  |||| |||
Db   429 AGGAAAATATATACTGACTTCATAGAAAAGGAAGCTCCAAAGAGATAAACATAGATTTT 488

Qy   625 CACACAAAGAAGTCATTACAAACAGCATCCACTCACTCACTCCACAGTTTTGTGATGCT 684
         ||  | ||| | |  |   ||| |||| ||||  |||  ||||||||||| ||| |||
Db   489 CAAACCAAAACTCTGATTGCCCAGAATATACAAGAAGCTACAAGTGGCTGCTTTACAACT 548

Qy   685 GCACAAAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTATACACGTTTTCTGAAATCT 744
         || ||||| | | ||||| |||| |||| ||| || |   |||| | ||| | ||| ||
Db   549 GCCCAGAAAGGGTATACAGCTTGATGGAGAACAACTCTTATCCTCGTTTCGTTCTTGGAGTCA 608

Qy   745 GACATCTATTTAGACTTG 762
         || ||| || ||||||||
Db   609 GAATTCTTACCAGGACTTG 626
```

RESULT 14
US-09-925-300-567
; Sequence 567, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 567
; LENGTH: 1364
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1362)
; OTHER INFORMATION: n equals a,t,g, or c

US-09-925-300-567

```
Query Match            7.0%;   Score 155.6;  DB 9;   Length 1364;
Best Local Similarity  63.2%;  Pred. No. 2.1e-20;
Matches  239;  Conservative   0;  Mismatches 139;  Indels   0;  Gaps   0;

Qy     385 TCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTGACAAACTGCTTTCCCATAGAGAT 444
           ||||| |||||||||||| || ||| ||  |||| |||| |||| | ||| |||||||||
Db     246 TCTCCTGAGGAAGCACAGCTGTGTCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 305

Qy     445 GGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAAGAAAATATTGAATTT 504
           |||| ||||||||| ||| |||| |||||||| |||||||||||| ||||||||||||||
Db     306 GGTCTTGCTGCATTCAGGGCTTTTTTAAAGTCGGAATTCTGTGAAGAAAATATTGAATTC 365

Qy     505 TGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAAAGCA 564
           |||||||||||||||| ||| ||||||||||||| ||| |||||||||| ||| ||||||
Db     366 TGGCTGGCCTGTGAAGACTTCAAAAAAACCAAATCACCCCAAAAGCTGTCCTCAAAAGCA 425

Qy     565 AAAGCAATATATGAGAAATTTATACGACTTGATGCCCCAAAAGAGGTTAACCTTGATTTT 624
           ||||||| ||| || ||||||||||| ||||||||| |||||||| ||| |||||||||
Db     426 AGGAAAATATATACTGACTTCATAGAAAAGGAAGCTCCAAAAGAGATAAACATAGATTTT 485

Qy     625 CACACAAAGAGAGTCATTACAAACAGCATCACCTCAACCTACCCTCCACAGTTTTGATGCT 684
           |||| |||| |||| |||| ||||| |||| || |||| ||| |||||||||||||||||
Db     486 CAAACCAAACTCTGATTGCCCAGAATATACAAGAGCTACAAGTGGCTGCTTACAACT 545

Qy     685 GCACAAAGCAGAGTGTATCAGCTCATGGAACAACAGAGTTATACACGTTTTCTGAAATCT 744
           ||||| ||||| ||||||||||||| |||| |||||| ||| |||||| ||||||||||
Db     546 GCCCAGAAAAGGGTATACAGCTTGATGGAGAGAACAACTCTTATCCTCGTTTCTTGGAGTCA 605

Qy     745 GACATCTATTAGACTTG 762
           |||||||||||||||||
Db     606 GAATTCTTACCAGGACTTG 623
```

RESULT 15
US-09-971-429B-32
; Sequence 32, Application US/09971429B
; Publication No. US20030175704A1
; GENERAL INFORMATION:
;   APPLICANT: Lasek, Amy K. W.
;   APPLICANT: Shyjan, Andrew W.
;   APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER
; FILE REFERENCE: PA-0040 US
; CURRENT APPLICATION NUMBER: US/09/971,429B
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/239,024
; PRIOR FILING DATE: 2000-04-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PERL Program
; SEQ ID NO 32
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030175704A1 989992.12
US-09-971-429B-32

```
Query Match            7.0%;   Score 155.6;  DB 10;  Length 1381;
Best Local Similarity  63.2%;  Pred. No. 2.1e-20;
Matches  239;  Conservative   0;  Mismatches 139;  Indels   0;  Gaps   0;

Qy     385 TCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTGACAAACTGCTTTCCCATAGAGAT 444
           ||||| |||||||||||| || ||| ||  |||| |||| |||| | ||| |||||||||
Db     262 TCTCCTGAGGAAGCACAGCTGTGTCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 321

Qy     445 GGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAAGAAAATATTGAATTT 504
           |||| ||||||||| ||| |||| |||||||| |||||||||||| ||||||||||||||
Db     322 GGTCTTGCTGCATTCAGGGCTTTTTTAAAGTCGGAATTCTGTGAAGAAAATATTGAATTC 381

Qy     505 TGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAAAGCA 564
           |||||||||||||||| ||| ||||||||||||| ||| |||||||||| ||| ||||||
Db     382 TGGCTGGCCTGTGAAGACTTCAAAAAAACCAAATCACCCCAAAAGCTGTCCTCAAAAGCA 441

Qy     565 AAAGCAATATATGAGAAATTTATACGACTTGATGCCCCAAAAGAGGTTAACCTTGATTTT 624
           ||||||| ||| || ||||||||||| ||||||||| |||||||| ||| |||||||||
Db     442 AGGAAAATATATACTGACTTCATAGAAAAGGAAGCTCCAAAAGAGATAAACATAGATTTT 501

Qy     625 CACACAAAGAGAGTCATTACAAACAGCATCACCTCAACCTACCCTCCACAGTTTTGATGCT 684
           |||| |||| |||| |||| ||||| |||| || |||| ||| |||||||||||||||||
Db     502 CAAACCAAACTCTGATTGCCCAGAATATACAAGAGCTACAAGTGGCTGCTTACAACT 561

Qy     685 GCACAAAGCAGAGTGTATCAGCTCATGGAACAACAGAGTTATACACGTTTTCTGAAATCT 744
           ||||| ||||| ||||||||||||| |||| |||||| ||| |||||| ||||||||||
Db     562 GCCCAGAAAAGGGTATACAGCTTGATGGAGAGAACAACTCTTATCCTCGTTTCTTGGAGTCA 621

Qy     745 GACATCTATTAGACTTG 762
           |||||||||||||||||
Db     622 GAATTCTTACCAGGACTTG 639
```

Search completed: August 20, 2004, 16:40:45
Job time : 1038.39 secs

```
                Copyright     GenCore version 5.1.6
                Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      August 20, 2004, 05:11:40 ; Search time 6012.98 Seconds
                                        (without alignments)
                                        11010.254 Million cell updates/sec

Title:            US-09-894-749-1
Perfect score:    2217
Sequence:         1 gaattcggcttcatcctaa........aaaaaaaaaggcggccgc 2217
Scoring table:    IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0

Searched:         27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters:   55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
          1: em_estba:*
          2: em_esthum:*
          3: em_estin:*
          4: em_estmu:*
          5: em_estov:*
          6: em_estpl:*
          7: em_estro:*
          8: em_htc:*
          9: gb_est1:*
         10: gb_est2:*
         11: gb_htc:*
         12: gb_htc:*
         13: gb_est4:*
         14: gb_est5:*
         15: em_estfun:*
         16: em_estom:*
         17: em_gss_hum:*
         18: em_gss_inv:*
         19: em_gss_pln:*
         20: em_gss_fun:*
         21: em_gss_rat:*
         22: em_gss_mam:*
         23: em_gss_mus:*
         24: em_gss_pro:*
         25: em_gss_rod:*
         26: em_gss_phg:*
         27: em_gss_vrl:*
         28: gb_gssl:*
         29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
c 1	998.2	45.0	1201	9	AL579846	AL579846 AL579846
c 2	918.2	41.4	1201	9	AL557903	AL557903 AL557903
3	872.2	39.3	923	13	BX349093	BX349093 BX349093
c 4	688.2	31.0	921	13	BX350311	BX350311 BX350311

No.		Score	%	Length	DB	ID	Description
5		680.4	30.7	703	14	CD642093	CD642093 AGENCOURT
6		655	29.5	708	29	AY420372	AY420372 Homo sapi
7		643	29.0	655	12	BG572311	BG572311 602594534
8		611	27.6	810	12	BG564257	BG564257 602586054
9		596	26.9	1860	11	AK036407	AK036407 Mus muscu
10		568	25.6	705	9	AV714060	AV714060 AV714060
11		565.4	25.5	639	9	AV716055	AV716055 AV716055
12		564.2	25.4	626	9	AV714044	AV714044 AV714044
13		557.6	25.2	723	14	CF181629	CF181629 818478 MA
14		550	24.8	661	14	CD466872	CD466872 LeukoN2_8
15		537	24.2	537	28	AQ317367	AQ317367 RPCI11-10
16		535.6	24.2	625	9	AV715708	AV715708 AV715708
17		519.8	23.4	652	14	CB553797	CB553797 MMSP0073
18		511.8	23.1	764	14	CD468328	CD468328 LeukoS3_2
19		507.6	22.9	697	12	BG253969	BG253969 60236648?
20		499.4	22.5	789	14	CD468976	CD468976 LeukoS3_8
21		486.8	22.0	701	9	AV712845	AV712845 AV712845_
22		465.2	21.0	589	29	AY420373	AY420373 Pan trogl
23		449.6	20.3	655	29	AY420374	AY420374 Mus muscu
24		441.6	19.9	448	28	AQ053459	AQ053459 RPCI11-50
25		419.6	18.9	538	13	BU698389	BU698389 LL2in1395
26		417.2	18.8	716	14	CD465070	CD465070 LeukoN1_1
27		409	18.4	411	9	AA195206	AA195206 zr34h12.r
28		404	18.2	431	28	AQ887601	AQ887601 HS_5557_A
29		395.8	17.9	473	14	W04905	W04905 za43510.r1
30	c	380.6	17.2	409	12	BM145481	BM145481 TCAAP1D53
31		380.2	17.1	700	13	BY750625	BY750625 BY750625
32		370	16.7	484	14	N98410	N98410 za69c05.r1
33		352.8	15.9	861	13	CF257072	CF257072 pha006_d0
34		348.6	15.7	618	13	BY722727	BY722727 BY722727
35		345.2	15.6	666	10	BB660526	BB660526 BB660526
36	c	334.4	15.1	472	14	N69945	N69945 za69c05.s1
37		322	14.5	772	13	BU285628	BU285628 603866709
38		306.4	13.8	616	10	BB629298	BB629298 BB629298
39		281.6	12.7	739	12	BI909559	BI909559 603070761
40		279.2	12.6	781	13	BU361362	BU361362 603788927
41		278.4	12.6	370	12	BM146419	BM146419 TCAAP1E47
42	c	277.6	12.5	674	14	CD465007	CD465007 LeukoN1_1
43		276	12.4	276	14	N59342	N59342 yz86e09.s1
44		273.6	12.3	706	14	CD468899	CD468899 LeukoS3_8
45		272.4	12.3	599	12	BG625427	BG625427 pgnic.pk0

ALIGNMENTS

```
RESULT 1
AL579846/c
LOCUS       AL579846      1201 bp    mRNA    linear   EST 01-JUN-2003
DEFINITION  AL579846 Homo sapiens cDNA clone CS0DJ003YJ10 3-PRIME, COT 10-NORMALIZED
            Homo sapiens cDNA clone CS0DJ003YJ10 3-PRIME, mRNA sequence.
ACCESSION   AL579846
VERSION     AL579846.2   GI:3131818126
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 16, 2001 this sequence version replaced gi:12945287.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 261.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DJ003DE05NP1&cluster=261.f Contact :
            Feng Liang Email : fliang@lifetech.com URL :
```

```
          http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
          Faraday Avenue Genoscope sequence ID : CS0DJ003DE05NP1.
FEATURES
   source  1..1201
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="CS0DJ003YJ10"
          /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
          /cell_line="JURKAT"
          /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
          10-NORMALIZED"
          /note="1st strand cDNA was primed with a NotI-oligo(dT)
          primer. Five prime end enriched, double-strand cDNA was
          digested with Not I and cloned into the Not I and EcoR V
          sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match     45.0%;    Score 998.2;  DB 9;   Length 1201;
Best Local Similarity  95.1%;  Pred. No. 1.2e-135;
Matches 1010;  Conservative  22;  Mismatches  29;  Indels  1;  Gaps  1;

Qy  1045 TTCTGTAAAACCTTCTATTTGATGTCATTCCATTTATAATCAGAAAAAAAACTTATTTCT 1104
         |||||||||||||| |||  ||| ||| ||||  |||||| |||||||||||||||||||
Db  1063 TWTTCTGTAAAACCCTYCTTTTTGTCATCCCATTTAT-ATCAGAAAAAAAACTTATTTCT 1005

Qy  1105 TAATCAAAAGGCAGTACAAAAAAAAGTAATAAATGTTTTATAAGATTGTAGAGTTAAGTAAA 1164
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1004 TATCMAAAAGGCAGTACAAAAAAAAAGTAATAAATGTTTTATAAGATTGTAGAGTTAAGTAAA 945

Qy  1165 AGTTAAGCTTTGCAAAGTGTCAAAAGTTCAAACAAAAGTCTAGTTGGGATTTTTTTACC 1224
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  944  AGTTAAGCTTTTGCAAAGTGTCAAAAGTTCAAACAAAAGTCTAGTTGGGATTTTTTTACC 885

Qy  1225 AAAGCAGCATAATGTGTTATAAAACATATAAACTCAGATATCCAAATGTTCAGAT 1284
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  884  AAAGCAGCATAATGTGTTATAAAACATATAAACTCAGATATCCAAATGTTCAGAT 825

Qy  1285 AGCATTTTCATAATGATGTTCCTTTTTTTTTTTTGGTAATAGTGTAGAAGTGATCTGGTTC 1344
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  824  AGCATTTTCATAATGATGTTCCTTTTTTTTTTTTGGTAATAGTGTAGAAGTGATCTGGTTC 765

Qy  1345 TTACAATGGGAGATGAAGAACATTTATTATTGGGTTACTACTAACCCTGTCCCAAGAATA 1404
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  764  TTACAATGGGAGATGAAGAACATTTATTATTGGGTTACTACTAACCCTGTCCCAAGAATA 705

Qy  1405 GTAATATCACCTCTAGTTATAAGCCAGCAACAGGAACTTTGTGAAGACACATTCATCTC 1464
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  704  GTAATATCACCTCTAGTTATAAGCCAGCRGCAACAGGAACTTTGTGAAGACACATTCATCTC 645

Qy  1465 TACAGAACTTCAGATTAAATAATATAATCTAGATTAATGACTAAGAAGAATAAGATCCACATTGA 1524
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  644  TACAGAACTTCAGATTAAATAATATAATCTAGATTAATGACTAAGAAGAATAAGATCCACATTGA 585

Qy  1525 ACTCATTCCTAAGTGAACATGAACATGGACGTACCCAGTTATACAAAGTACTTCTGTTGGTCACAG 1584
         |||||||||||||||||| || |||||||||||||||||||||||||||||||||||||||::|
Db  584  ACTCATTCCTAAGTGAACATGAACATGGACGTACCCAGTTATACAAAGTACTTCTGTTGGTCACMG 525

Qy  1585 AAACATGACCAGATTTGCATATCTCCAGGTAGGGAACTAAGTAGACTACCTTATCCACCG 1644
         ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  524  AAACATGACCRGATTTGCATATCTCCAGGTAGGGAACTAAGTAGACTACCTTATCCACCG 465

Qy  1645 GCTAAGAAACTTCGCTACTAAACTATTAGGCCATCAAATGGCTTGAATAAAAACCAGAGAA 1704
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  464  GCTAAGAAACCCTTGCTACTAAACTATTAGGCCATCAAATGGCTTGAATAAAAACCAGAGAA 405

Qy  1705 GGTTTTTCCCAGGACGTCTCATGTTTGGCCCTTTAGAATTGGGGTAGAAATCAGAAATGA 1764
         ||||  |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  404  GGTTTTYCCCAGGACGCCTTCATGTTTGGCCCTTTAGAATTGGGGTGGTRGAAATCMGAAATGA 345

Qy  1765 GATGAGGGGAAGAAGAGCAAGGAGTCAAGGAGTCCCTAGCCCCTAGCCCCTAGCCCATTGGT 1824
         ||||||||||||||||||||| ||||||||||||| :||| |||| :|||||||||:|||
Db  344  GATGAGGGGAAGAAGAGCAAGGAGTCTAAGGCCCTAGCGMTTTGGGCACYCTGCCCCCATTGGT 285
```

```
Qy  1825 TCATATTCAGAAAGTGTTATTCATTCATTGATTATATTCTGTTGTAAGCAAATCTCCTTAAGTA 1884
         ||||||| |||||| ||||||||||||||||||| |||||| ||||||||||| |||||||||||
Db  284  TCATATTCMGAAAGTGTTWTTCATTGATTATWTTCTTGTTGTTAAGCAAATCYCCTTAAGTA 225

Qy  1885 ATTATTATTCAAATAAGATATACTCATACATCATATATGTCACTGTTTTAAAGAGATATT 1944
         ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  224  ATTATTATTCAAATAAGRTTATACYCATACAMTCTATATGTCCCGTTTTAAAGAGATATT 165

Qy  1945 TAATTTTTAATGTGTGTTACATGGGTCGTGTAAAATATTTGTATTTAAAAATGCCATGCATTA 2004
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  164  TAATTTTTAATGTGTGTTACATGGGTCTGTAAAATATTTGTATTTAAAAATGCCATGCATTA 105

Qy  2005 GGCTTTGGAAATTAATGTTAGTTGAAAATGTAAAATGTAAAACTTTAGATCATTGTAG 2064
         |||||||||||||||||||||||||||||||||||||||||||||||||||| :|||||
Db  104  GGCTTTGGAAATTTAATGTTAGTTGAAAATGTAAAATGTAAAAMCTTAGATCCTTTGTAG 45

Qy  2065 TAATAAATATTTTAACTTCATTCATACAGTTAAGTTATCT 2106
         ||||| |||||||  ||||| ||||||||::
Db  44   TAWTTAATATTTTTNACTTCATTTATACCGTTAANNNNNBB 3
```

```
RESULT 2
AL557903
LOCUS       AL557903               1201 bp    mRNA       linear   EST 31-MAY-2003
DEFINITION  AL557903 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
            Homo sapiens cDNA clone CS0DJ003YJ10 5-PRIME, mRNA sequence.
ACCESSION   AL557903
VERSION     AL557903.2  GI:31279703
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12901962.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 261.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DJ003DE05QP1&cluster=261.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DJ003DE05QP1.
FEATURES
   source  1..1201
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="CS0DJ003YJ10"
          /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
          /cell_line="JURKAT"
          /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
          10-NORMALIZED"
          /note="1st strand cDNA was primed with a NotI-oligo(dT)
          primer. Five prime end enriched, double-strand cDNA was
          digested with Not I and cloned into the Not I and EcoR V
          sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match     41.4%;    Score 918.2;  DB 9;   Length 1201;
Best Local Similarity  97.0%;  Pred. No. 4.9e-124;
Matches 938;  Conservative  7;  Mismatches  20;  Indels  2;  Gaps  1;

Qy  61   ATAACTTTTATTCTACTATGTATGTATGTGGAAGTAGTATTAATAAATGAACTAGGGAAG 120
         ||||||||||||||||||||||||||||||| ::||||||||| :||||||
Db  344  GATGAGGGGAAGAAGAGCAAGGAGTCAAGGAGTCCTAGCGMTTTGGGCAYCTGCCCCATTGGT 285
```

```
VERSION      BX349093.1  GI:30375347
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 923)
  AUTHORS    Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE      Full-length cDNA libraries and normalization
  JOURNAL    Unpublished (2001)
COMMENT      Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             Library was constructed by Life Technologies, a division of
             Invitrogen. This sequence belongs to sequence cluster 261.f For
             more information about this cluster, see
             http://www.genoscope.cns.fr/
             cgi-bin/cluster.cgi?seq=CS0BAG041ZE09_CS03909_1&cluster=261.f.
             Contact: Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
             Faraday Avenue Genoscope sequence ID : CS0BAG041ZE09_CS03909_1.
FEATURES             Location/Qualifiers
     source          1..923
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ003YJ10"
                     /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                     /cell_line="JURKAT"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
                     10-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match         39.3%;  Score 872.2; DB 13;  Length 923;
Best Local Similarity   97.8%;  Pred. No. 2.6e-117;
Matches  902;  Conservative  0;  Mismatches  18;  Indels  2;  Gaps  2;

Qy   865 TAAAGAAATTGATTTTGCTCATTTTTATGACAAACTTATACATCTGCTTCTAACATATC 924
             ||||||| |||||| || ||||||||||||||||||||||| ||||||||||||||||
Db     2 TTAAAAAATAGATTTGCTTCATTTTCATTTTTATGACAAACTATACATCTGCTTCTAACATATC 61

Qy   925 GCATGTTTATGTTAAGATTTGGTCCCATCCTTTAAACTGAAATATGTCATGTGAAATTAT 984
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    62 GCATGTTTATGTTAAGATTTGGTCCCATCCTTTAAACTGAAAT-TGTCATGTGAAATTAT 120

Qy   985 TTTAAAAATGTAAAACAAAACTTTCTGCTACAAAATACATACAGTATCTGCCAGTATA 1044
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TTTAAAAATGTAAAACAAAACTTTCTGCTAACAAAATACATACAGTATCTGCCAGTATA 180

Qy  1045 TTCTGTAAAACCTTCATTTGATGTCATTCCATTTATTTATAATCAGAAAAAAACTTATTTCT 1104
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TTCTGTAAAACCTTCATTTGATGTCATTCCATTTATTTATAATCAGAAAAAAACTTATTTCT 240

Qy  1105 TAATCAAAAGGCAGTACAAAACTTTCTGCTACAAAAAAGTAATAATGTTTTATAAGAGTTAAGTAAA 1164
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TAATCAAAAGGCAGTACAAAACTTTCTGCTACAAAAAAGTAATAATGTTTTATAAGAGTTAAGTAAA 300

Qy  1165 AGTTAAGCTTTTGCAAAGTGTCAAAAGTGTCAAAAGTCTAGTTGGGATTTTTTTACC 1224
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGTTAAGCTTTTGCAAAGTGTCAAAAGTGTCAAAAGTCTAGTTGGGATTTTTTTACC 360

Qy  1225 AAAGCAGCATAATATGTGTTATATAAACATAATAAACATAATAAACATACTCAGATATCCAAATGTTCAGAT 1284
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AAAGCAGCATAATATGTGTTATATAAACATAATAAACATAATAAACATACTCAGATATCCAAATGTTCAGAT 420

Qy  1285 AGCCATTTTCATAATGAATGTTCTCTTTTTTTTGGTAATAGTGTGAAGTGATCTCGGTTC 1344
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGCCATTTTCATAATGAATGTTCTCTTTTTTTTGGTAATAGTGTGAAGTGATCTCGGTTC 480
```

```
Db   110 ATAACTTTTTATTCTACTATGTATGTATGGAAWAGTATTAAWAAATGAACTAGGGAAG 169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121 GATGTAATAAATTAGACATCTCCTCATTTAGAGAGAAGATGAAAACAACATTGCTTTTC 180

Db   170 GATGTAATAAATTAGACATCTCTTCATTTAGAGAGAAGATGAAAACAACATTGCTTTTC 229
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   181 TTTTCTCAAATAAATATGTGTGAATCAAAAGAAAAAACTTTTTTCAAGTAATACATGGT 240

Db   230 TTTTCTCAAATAAATATGTGTGAATCAAAAGAAAAAACTTTTTTCAAGTTAATACATGGT 289
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   241 TCAGGAAAAGAAGAACAAGCAAAGAAGCAAATCAGAGCTAAGGAAAAAAGAAAATAGA 300

Db   290 TCAGGAAAAGAAGAAACAGCAAAGAAGCAAATCAGAGCTAAGGAAAAAAGAAAATAGA 349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   301 CTAAGTCTTCTTGTGCAGAAACCTGAGTTCATGAAGACACCCGGCTCCAGTAGATCTGGG 360

Db   350 CTAAGTCTTCTTGTGCAGAAACCTGAGTTCATGAAGACACCCGGCCTCCAGTAGATCTCGG 409
             ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Qy   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAGAGGCAGTGAAATGGGTGAATCATTT 420

Db   410 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAGAGGCAGTGAAATGGGTGAATCATTT 469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   421 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTCTTAAAACTGAA 480

Db   470 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTCTTAAAACTGAA 529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   481 TTCAGTGAAGAAAAATATTGAATTTTTTGGATAGCCTGTGAAGATTTCAAGAAAGCAAGGGA 540

Db   530 TTCAGTGAAGAAAAATATTGAATTTTTTGGATAGCCTGTGAAGATTTCAAGAAAGCAAGGGA 589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541 CCTCAACAAATTCACCTAAAGCAAAGCAAATATATGAGAAATTTATACAGACTGATGCC 600

Db   590 CCTCAACAAATTCACCTAAAGCAAAGCAAATATATGAGAAATTTATACAGACTGATGCC 649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   601 CCAAAAGAGGTAACCTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAA 660

Db   650 CCAAAAGAGGTAACCTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAA 709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   661 CCTACCCCCACAGTTTGATGCTGCACAAAGCAGAGAGTGTATCAGCTTCATGGAACAGAC 720

Db   710 CCTACCCCTCCACAGTTTGATGCTGCACAAAGCAGAGAGTGTATCAGCTTCATGGAACAGAC 769
             ||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Qy   721 AGTTATACACGTTTTCTGAAATCTGACATCTATTTAGACTTGAGAAGGAAGAGACCTCAG 780

Db   770 AGTTATACACGTTTTCTGAAATCTGACATCTATTTAGACTTGAGAAGGAAGAGACCTCAG 829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   781 AGACCAACAAATCTTAGGAGACGATCACGTCATTTACCTGCAATGAATTCCAAGATGTA 840

Db   830 AGACCAACAAATCTTAGGAGACGATCACGTCATTTACCTGCAATGAATTCCAAGATGTA 889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   841 CAATCAGATGTGTGCCATTTGGTTATAAAGAAAATTGATTTTGCTCATTTTTATGACAAAC 900

Db   890 CAATCAGATGTGTGCCATTTGGTTATAAAGAAAATTGATTTTGCTCATTTTTATGACAAAC 949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   901 TTCTGTAAAACCTTCTTCTTCAACATATCGCATGTTTAGAGATTTGGTCCCATCCTTTAAA 960

Db   950 TTTATACACGTTTCTCTAACATATCGCATGTTTATGTTTAGAGATTTGGTCC--ATCCTTAAA 1007
             |||:||:||||||:||||||||||||||||||||||||||||||||  |||||||
Qy   961 CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACTTTCTGCTACAAA 1020

Db  1008 CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAAMAAAACTTTCTGCTAACAAA 1067
             |||||||||||||||||||||||||:||:      |||||||||||:|
Qy  1021 ATACATA 1027
             ||:::|
Db  1068 ATAMRGA 1074
```

```
RESULT 3
BX349093
LOCUS       BX349093         923 bp    mRNA    linear    EST 05-MAY-2003
DEFINITION  BX349093 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
            Homo sapiens cDNA clone CS0DJ003YJ10 5-PRIME, mRNA sequence.
ACCESSION   BX349093
```

primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match   31.0%; Score 688.2; DB 13; Length 921;
Best Local Similarity 94.0%; Pred. No. 1.2e-90;
Matches 758; Conservative 0; Mismatches 44; Indels 4; Gaps 4;

```
Qy   392  AAGAGGCAGTGAAATGGGTGAATCATTGCAAACTGCTTCCCATAGAGATGGACTAG 451
               ||||||||||||||||| ||||||||||||||||||  |||  | ||||||||||
Db   804  AAAAAAGAGAAAAAGGGGGGGCAACCATTCTAAACCCCGGTTTTCCAAAAACAGGGACTAC 745

Qy   452  AGGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAAGAAAATATTGGAATTTGGATAG 511
               ||||||||||||||||||||||||||||||||||||||||| | |  ||||||||||
Db   744  AGCTTTTCACCCAGATTCTTAACTTCAATTCAGTT-AACAAAATAATGATTTTTGGATAG 686

Qy   512  CCTGTGAAGATTTCAAGAAAGC-AAGGGACCTCAACAA-TTCACCTTAAAGCAAAAGC 569
               ||||||||||||||||||||| ||||||||||||||  ||||||||||||||||||
Db   685  CCTGTGAAGATTTCAAGAAAGCAAGGGACCTCAACAAANTCACCTTAAAGCAAAAGC 626

Qy   570  AATATATGAGAAATTTATACAGACAGCATCATCAACCTGATTTTCACAC 629
               ||||| |||||| |||||||||||||||||||||||||||| |||||||||
Db   625  ATATATGAGANATTTATACAGACAGCTGATGCCTCAAAAAGAGGTTAACCTT-ATTTTCACAC 567

Qy   630  AAAAGAAGTCATTACAACAAACAGCATCAACCTACCCTCCAACAAAATCTTAGGAGACGATCACG 689
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   566  AAAAGAAGTCATTACAACAAACAGCATCAACCTACCCTCCAACAAAATCTTAGGAGACGATCACG 507

Qy   690  AAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTATACACGTTTTCTGAAATCTGACAT 749
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   506  AAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTATACACGTTTTCTGAAATCTGACAT 447

Qy   750  CTATTTAGACTTGATGGAAGGAAGGAAGGAGACCTCAGAGACCAACAAATCTTAGGAGACGATCACG 809
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   446  CTATTTAGACTTGATGGAAGGAAGGAAGGAGACCTCAGAGACCAACAAATCTTAGGAGACGATCACG 387

Qy   810  CTCATTTACCTGCAATGAATTCCAAGATGTACAATCAGATGTTGCCATTGGTTATAAAG 869
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   386  CTCATTTACCTGCAATGAATTCCAAGATGTACAATCAGATGTTGCCATTGGTTATAAAG 327

Qy   870  AAAATTGATTTTGCTCATTTTATGACAAACTTATACATCTGCTTCTAACATATCGCATG 929
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   326  AAAATTGATTTTGCTCATTTTATGACAAACTTATACATCTGCTTCTAACATATCGCATG 267

Qy   930  TTTATGTTAAGATTGGTCCCATCCTTAAACTGAAAATATGTCATGTGAAATATTTTTAA 989
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   266  TTTATGTTAAGATTGGTCCCATCCTTAAACTGAAAATATGTCATGTGAAATATTTTTAA 207

Qy   990  AAATGTAAAAACAAAACTTTCTGCTAACAAAATACACAGTATCTGCCAGTATATTCTG 1049
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   206  AAATGTAAAAACAAAACTTTCTGCTAACAAAATACACAGTATCTGCCAGTATATTCTG 147

Qy  1050  TAAAACTTCTCATTTGATGTCATTCCATTTATAATCAGAAAAAAACTTATTTCTTAATC 1109
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   146  TAAAACTTCTCATTTGATGTCATTCCATTTATAATCAGAAAAAAACTTATTTCTTAATC 87

Qy  1110  AAAAGGCAGTACAAAAAAGTAATAAATGTTTATAAGAGTTAAGTAAAAGTTA 1169
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    86  AAAAGGCAGTACAAAAAAGTAATAAATGTTTATAAGAGTTAAGTAAAAGTTA 27

Qy  1170  AGCTTTTGCAAAGTGTCAAAAGTTC 1195
               |||||||||||||||||||||||||
Db    26  AGCTTTTGCAAAGTGTCAAAAGTTC 1
```

RESULT 5
CD642093
LOCUS   CD642093   703 bp   mRNA   linear   EST 17-JUN-2003
DEFINITION   AGENCOURT 14536786 NIH_MGC_191 Homo sapiens cDNA clone
             IMAGE:30418254 5', mRNA sequence.
ACCESSION   CD642093
VERSION   CD642093.1 GI:31811624

```
Qy  1345  TTACAATGGGAGATGAAGAACATTTATTATTGGGTTACTACTAACCCTGTCCCAAGAATA 1404
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TTACAATGGGAGATGAAGAACATTTATTATTGGGTTACTACTAACCCTGTCCCAAGAATA 540

Qy  1405  GTAATATCACCTCTAGTTATAAGCCAGCAACAGGAACTTTTGTGAAGACACATTCATCTC 1464
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  GTAATATCACCTCTAGTTATAAGCCAGCAACAGGAACTTTTGTGAAGACACATTCATCTC 600

Qy  1465  TACAGAACTTCAGATTAATAATATCTAGATTAATGACTAGAATTAGATCCACATTTGA 1524
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  TACAGAACTTCAGATTAATAATATCTAGATTAATGACTAGAATTAGATCCACATTTGA 660

Qy  1525  ACTCATTCCTAAGTGAACATGAACGTACCCAGTTATACAAAGTACTTCTGTTGGTCACAG 1584
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  ACTCATTCCTAAGTGAACATGAACGTACCCAGTTATACAAAGTACTTCTGTTGGTCACAG 720

Qy  1585  AAACATGACCAGATTTGCATATCTCCAGGTAGGGACTAAGTAGACTACCTTATCACCG 1644
               ||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db   721  AAACATGACCAGATTTGCATATCNTCCAGGTAGGGACTAAGTAGACTACCTTATCACCG 780

Qy  1645  GCTAAGAAAACTGCTACTAAACTATTAGGCCATCAATGGCTTGAATAAAAACCAGAGAA 1704
               ||||||||||||||||||||||  |||||||||||||||||||| ||||||||||||||
Db   781  NCTAAGAAAACTGCTACTAAACTATTAGGCCATCAATGGCTTGAATAAAAACCAGAAAA 840

Qy  1705  GGTTTTTCCAGGACGTCTCATG-TTTGGCCCTTTAGAATTGGGGTAGAAATCAGAAATG 1763
               ||||||||||||||||||||   ||||||||||||||||||  |||||||||||||||
Db   841  GNTTTTTTCCAGAACGTCTCATGTTTGCCCTTTAGAATTGNNGTAGAAATCAGAAATG 900

Qy  1764  AGATGAGGGGAAGAAGCAAGGA 1785
               ||||||||||||||||||||||
Db   901  AGATGAGGGGAAGAAGCAAGGA 922
```

RESULT 4
BX350311/c
LOCUS   BX350311   921 bp   mRNA   linear   EST 05-MAY-2003
DEFINITION   BX350311 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
             Homo sapiens cDNA clone CS0DJ003YJ10 3-PRIME, mRNA sequence.
ACCESSION   BX350311
VERSION   BX350311.1 GI:30373474
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 921)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
             Library was constructed by Life Technologies, a division of
             Invitrogen. This sequence belongs to sequence cluster 261.f For
             more information about this cluster, see
             http://www.genoscope.cns.fr/
             cgi-bin/cluster.cgi?seq=CS0BAI032ZC02_CS02987_1&cluster=261.f.
             Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fullength.invitrogen.com/ InVitroGen Corporation 1600
             Faraday Avenue Genoscope sequence ID : CS0BAI032ZC02_CS02987_1.
FEATURES   Location/Qualifiers
     source   1..921
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ003YJ10"
                     /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                     /cell_line="JURKAT"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
                     10-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)

```
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 703)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Daniela S. Gerhard, Ph.D.
           Office of Cancer Genomics
           National Cancer Institute / NIH
           Bldg. 31 Rm10A07 Bethesda, MD 20892
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Narayan Bhat
           cDNA Library Preparation: CLONTECH Laboratories, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: NDCM220  row: p  column: 07
FEATURES   High quality sequence stop: 616.
  source   1..703
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /clone="IMAGE:30418254"
           /tissue_type="Pooled"
           /lab_host="DH10B (T1 phage-resistant)"
           /clone_lib="NIH_MGC_191"
           /note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc);
           Site_2: SfiI (ggccgcctcggcc); Library is oligo-dT primed
           and directionally cloned. PBMC - Peripheral Blood
           Mononuclear Cells. RNA was pooled from 3/6hour stimulation
           with PMA adn Ionomycin. 5' and 3' adaptors were used in
           cloning as follows: 5' adaptor sequence:
           5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
           5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
           C, or G and N = A, C, G, or T). Average insert size 1.69
           kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
           by PCR. This library was enriched for full-length clones
           and was constructed by Clontech Laboratories (Palo Alto,
           CA). Note: this is a NIH_MGC Library."
ORIGIN

Query Match      30.7%;  Score 680.4;  DB 14;  Length 703;
Best Local Similarity   98.1%;  Pred. No. 1.9e-89;
Matches  687;  Conservative    0;  Mismatches   13;  Indels   0;  Gaps    0;

Qy    973 ATGTGAAATTATTTAAAAATGTAAAACAACAAAACTTCTGCTAACAAAATAACATACAGTA 1032
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      3 ATGTGAAATTATTTAAAAATGTAAAACAACAAAACTTCTGCTAACAAAATAACATACAGTA   62

Qy   1033 TCTGCCAGTATATTCTGTAAAACCTTCATTGATGTCATTCATTCAGAAAAA 1092
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 TCTGCCAGTATATTCTGTAAAACCTTCATTGATGTCATTCATTCAGAAAAA  122

Qy   1093 AAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAGTAATAATGTTTTATAAGATTGTA 1152
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 AAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAGTAATAATGTTTTATAAGATTGTA  182

Qy   1153 GAGTTAAGTAAAAGTTAAGCTTTTGCAAAGTTCAAACAAAGTTCAAACAAAGTCTAGTTG 1212
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 GAGTTAAGTAAAAGTTAAGCTTTTGCAAAGTTGTCAAAGTTCAAACAAAGTCTAGTTG  242

Qy   1213 GGATTTTTTACCAAAGCAGCATAATATGTGTTATATAAACATAATAATACTCAGATATCC 1272
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 GGATTTTTTACCAAAGCAGCATAATATGTGTTATATAAACATAATAATACTCAGATATCC  302

Qy   1273 AAATGTTCAGATAGCATTTTCATAATGAATGTTCCTCATTTATTTGGTAATAGTGTAGAA 1332
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    303 AAATGTTCAGATAGCATTTTCATAATGAATGTTCCTCATTTTTTGGTAATAGTGTAGAA  362
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1333 GTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTTACTACTAACCCT 1392
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   '363 GTGATCTGGTTCTTACAATGGGAGATGAAGAACATTTATTATTGGGTTACTACTAACCCT  422

Qy   1393 GTCCCAAGAATAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAACTTTGTGAAGA 1452
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 GTCCCAAGAATAGTAATATCACCTCTAGTTATAAGCCAGCAACAGGAACTTTGTGAAGA  482

Qy   1453 CACATTCATTCTCACAGAACTTCAGATTAAATAATAATCTAGATTAATGACTGAGAATAAG 1512
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    483 CACATTCATTCTCACAGAACTTCAGATTAAATAATAATCTAGATTAATGACTGAGAATAAG  542

Qy   1513 ATCCACATTTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTATACAAAGTACTTC 1572
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 ATCCACATTTGAACTCATTCCTAAGTGAACATGGACGTACCCAGTTATACAAAGTACTTC  602

Qy   1573 TGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCAGTAGGGAACTAAGTAGACT 1632
          ||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db    603 TGTTGGTCACAGAAACATGACCAGATTTGCATATCTCCCAGTAGGNAAACTAGTAGACT  662

Qy   1633 ACCTTATCACCGGCTAAGAAAACTTGCTACTAAACTATTA 1672
          ||||||||||||||||||||||||||| |||||||||||
Db    663 ACCTTTATCACCGGCTAAGAAAAACTTGCTACTAANCTATTA  702
```

```
RESULT 6
AY420372
LOCUS      AY420372                 708 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION Homo sapiens HCM7202 gene, VIRTUAL TRANSCRIPT, partial sequence,
           genomic survey sequence.
ACCESSION  AY420372
VERSION    AY420372.1  GI:39776329
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 708)
AUTHORS    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
           Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
           Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
           Adams,M.D. and Cargill,M.
TITLE      Inferring nonneutral evolution from human-chimp-mouse orthologous
           gene trios
JOURNAL    Science 302 (5652), 1960-1963 (2003)
PUBMED     14671302
REFERENCE  2 (bases 1 to 708)
AUTHORS    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
           Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
           Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
           Adams,M.D. and Cargill,M.
TITLE      Direct Submission
JOURNAL    Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
           Rockville, MD 20850, USA
COMMENT    This sequence was made by sequencing genomic exons and ordering them
           based on alignment.
FEATURES   Location/Qualifiers
  source   1..708
           /organism="Homo sapiens"
           /mol_type="genomic DNA"
           /db_xref="taxon:9606"
  gene     <1..>708
           /locus_tag="HCM7202"
ORIGIN

Query Match      29.5%;  Score 655;  DB 29;  Length 708;
Best Local Similarity   92.5%;  Pred. No. 9.3e-86;
Matches  655;  Conservative    0;  Mismatches    0;  Indels   0;  Gaps    0;

Qy    160 ATGGAAACAACATTGCTTTCTTTCTCAAATAAATAAATGTGTGAATCAAATAAAAGAAAACT 219
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1 ATGGAAACACATGCTTTTCTTTTCTCAAATAAATATGTGTGAATCAAAGAAAAAACT 60

Db    220 TTTTTCAAGTTAATACATGGTTCAGGAAAGAAGAAGAAGCAAAGAAGCCAAAATCAGA 279
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 TTTTTCAAGTTAATACATGGTTCAGGAAAGAAGAAGAAGCAAAGAAGCCAAAATCAGA 120

Qy    280 GCTAAGGAAAAAGAAAATAGACTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAGAC 339
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCTAAGGAAAAAGAAAATAGACTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAGAC 180

Qy    340 ACCCGCTCCAGTAGATCTGGGCACTTGGCCAAAGAAACAAAGAAGTCTCCCCTGAAGAGGCA 399
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCCGCTCCAGTAGATCTGGGCACTTGGCCAAAGAAACAAAGAAGTCTCCCCTGAAGAGGCA 240

Qy    400 GTGAAATGGGGTGAATCATTGACAAACTGCTTCCCATAGAGATGGACTAGGAGGCTTTT 459
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGAAATGGGGTGAATCATTTGACAAACTGCTTTCCCATAGAGATGGACTAGGAGGGTTTT 300

Qy    460 ACCAGATTTCTAAAACTGAATTCAGTGAAGAAAAATATTGAATTTGGATAGCCTGTGAA 519
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACCAGATTTCTAAAACTGAATTCAGTGAAGAAAAATATTGAATTTGGATAGCCTGTGAA 360

Qy    520 GATTTCAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAAAGCAAAAGCAATATATGAG 579
          ||||||||||||||||||||||||||||||||||||||||||NNNNNNNNNNN
Db    361 GATTTCAAGAAAAGCAAGGGACCTCAACAAATTCACCNNNNNNNNNNNNNNNNNN 420

Qy    580 AAATTTATACGACTGATGCCCCAAAGAGAGGTAACCTGATTTCACACAAAGAGAGTC 639
          ||||||||||||||||||||||NNNNNNNNNNNN
Db    421 NNNNNNNNNNNNNNNNNNNNNNNGTTAACCTGATTTCACACAAAGAGAGTC 480

Qy    640 ATTACAAACAGCATCACTCAACCTACCCTCCACAGTTTGATGCTGCCACAGAGTG 699
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATTACAAACAGCATCACTCAACCTACCCTCCACAGTTTGATGCTGCCACAGAGTG 540

Qy    700 TATCAGCTCATGGAACAACAGCAGTTATACACGTTTCTGAAAATCGACATCATTAGAC 759
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TATCAGCTCATGGAACAACAGCAGTTATACACGTTTCTGAAAATCGACATCATTAGAC 600

Qy    760 TTGATGGAAGGAAGACCTCAGAGACCAACAAAAATCTTAGGAGACGATCACGCTCATTACC 819
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTGATGGAAGGAAGACCTCAGAGACCAACAAAAATCTTAGGAGACGATCACGCTCATTACC 660

Qy    820 TGCAATGAATTCCAAGATGTACAATCAGATGTTGCCATTTGGTTATAA 867
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TGCAATGAATTCCAAGATGTACAATCAGATGTTGCCATTTGGTTATAA 708
```

```
RESULT 7
BG572311
LOCUS       BG572311                        655 bp    mRNA        linear   EST 10-APR-2001
DEFINITION  602594534F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4721624 5',
            mRNA sequence.
ACCESSION   BG572311
VERSION     BG572311.1  GI:13579964
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC).
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1578 row: d  column: 09
```

FEATURES   Location/Qualifiers
     source   1..655
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4721624"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_79"
                     /note="Organ: placenta; Vector: pDNR-LIB (Clontech);
                     Site 1: SfiI (ggccatatggc); Site 2: SfiI
                     (ggccattatggcc); 5' and 3' adaptors were used in cloning
                     as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'
                     and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C, G, or T). Average insert size 1.3
                     kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."

High quality sequence stop: 655.

ORIGIN

```
Query Match             29.0%;  Score 643;  DB 12;  Length 655;
Best Local Similarity   99.8%;  Pred. No. 5.4e-84;
Matches  654;           Conservative    0;  Mismatches    0;  Indels    0;  Gaps    1;

Qy   1242 GTTATATAAACATAATAATATCCAGATATCCAAATGTTCAGATAGCATTTTCATAATGA 1301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTTATATAAACATAATAATATCCAGATATCCAAATGTTCAGATAGCATTTTCATAATGA 60

Qy   1302 ATGTTCTCTTTTTTTTGGTAATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAA 1361
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATGTTCTCTTTTTTTTGGTAATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAA 120

Qy   1362 GAACATTTATTATTGGGTTACTACTACCCTGTCCCAAGAATAGTAATATCACCTCTAGT 1421
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GAACATTTATTATTGGGTTACTACTACCCTGTCCCAAGAATAGTAATATCACCTCTAGT 180

Qy   1422 TATAAGCCAGCAACAGGAACTTTGTGAAGAATAAGATCCACATTCACCATTCCTAAGTGAA 1481
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TATAAGCCAGCAACAGGAACTTTGTGAAGAATAAGATCCACATTCACCATTCCTAAGTGAA 240

Qy   1482 AATATAATCTAGATAATGACTGAGAATAAGACTACTTCTGTTGGTCACAGAAACATGACCAGATTTT 1541
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AATATAATCTAGATAATGACTGAGAATAAGACTACTTCTGTTGGTCACAGAAACATGACCAGATTTT 300

Qy   1542 CATGGACGTACCCAGTTATACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTT 1601
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CATGGACGTACCCAGTTATACAAAGTACTTCTGTTGGTCACAGAAACATGACCAGATTTT 360

Qy   1602 GCATATCTCCAGGTAGGAACTAAGTAGACTAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTA 1661
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCATATCTCCAGGTAGGAACTAAGTAGACTAGTAGACTACCTTATCACCGGCTAAGAAAACTTGCTA 420

Qy   1662 CTTAAACTATTAGGCCATCAATGGCTTGAATAAAACCAGAGAGAAGGTTTTCCCAGGACGT 1721
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTTAAACTATTAGGCCATCAATGGCTTGAATAAAACCAGAGAGAAGGTTTTCCCAGGACGT 480

Qy   1722 CTCATGTTGGCCCTTAGAATTGGGGTAGAATTGGGGTAGAAATGAGATGAGGGGAAGAA-GC 1780
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTCATGTTGGCCCTTAGAATTGGGGTAGAATTGGGGTAGAAATGAGATGAGGGGAAGAACGC 540

Qy   1781 AAGGAGTCTAAGGACCCCTAGGCATTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTG 1840
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAGGAGTCTAAGGACCCCTAGGCATTGGGCATCTGCCACATTGGTTCATATTCAGAAAGTG 600

Qy   1841 TTATCTCCATTGATTATATTCTTGTGTTAAGCAAATCTCCTTAAGTAATAATTATTATTCA 1895
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTATCTCCATTGATTATATTCTTGTGTTAAGCAAATCTCCTTAAGTAATAATTATTATTCA 655
```

RESULT 8
BG564257

```
Qy    420  TGACAAACTGCTTCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGA  479
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431  TGACAAACTGCTTTCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGA  490

Qy    480  ATTCAGTGAAGAAAATATTGAATTTTGGATAGCCTGTGAGATTTCAAGAAAAGCAAGGG  539
           |||| |||||||||||| ||||||||| ||||||||||||||||||||||||||||||
Db    491  ATTCAGTGAAGAAAATATGCGAATTTCGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAACCAAGGG  550

Qy    540  ACCTCACAAATTCACCTTAAAGCAAAAGCAATATATGAG-AAATTTATACAGACTGATG  598
           |||||||||||||||||||||||||||||||||| |||||   |||||||||||||||
Db    551  ACCTCACAAATTCACCTTAAAGCAAAAGCAATATATGAGCAAATTTATACAGACTGATG  610

Qy    599  CCCCAAAAGAGGTTAACCTTGATTT--CACACAAAAGAGTCATTACAAACAGCATCAC  656
           |||||||||||||||||||||||||   |||||| ||||||||||||||||||||||
Db    611  CCCCAAAAGAGGTTAACCTTGATTTTTCACACCCAAAAGAGTCATTACAAACAGCATCAC  670

Qy    657  TCAACCTACCCTCCACAGTTTGATCGTCGACAAA---GCAGAGTGTATCAGCTCATGGA  713
           ||||| ||||| ||||||||||||| |||  ||||   |||||||||||||||| |||||
Db    671  TCAAGCTACCTCCACATGTTTGATGCTGCGCACAAAAGCAAGAGTGTATCAGCTCCATGG  730

Qy    714  ACAAGACAGTATACACGTTTCTGAAATCTGCACATCTATTTGACTTGATGGAAGGAAG  773
           | ||||||||| ||||| ||||||||||||||| ||||| |||| |||||||||||||
Db    731  AACAAGACAGTATACACGTCCTCCTGAAAATCTGACATCATCTATTTAGAACTGGCTGGAAGGCAG  790

Qy    774  ACCTCAGAGACCAACAA  790
           | |||  |||| ||||||
Db    791  ACTTCCGAGACCACCAAA  807
```

RESULT 9
AK036407

LOCUS   AK036407   1860 bp   mRNA   linear   HTC 19-SEP-2003
DEFINITION   Mus musculus adult male bone cDNA, RIKEN full-length enriched
             library, clone:9830006J10 product:regulator of G-protein signaling
             18, full insert sequence.
ACCESSION   AK036407
VERSION   AK036407.1   GI:26331373
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
  PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
             Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
             prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
  PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
             Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
             Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
             Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
             Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
             Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
             Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system-384 multiformat
             sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530013
  PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
             FANTOM Consortium.
  TITLE   Functional annotation of a full-length mouse cDNA collection
```

---

```
LOCUS        BG564257                 810 bp    mRNA    linear    EST 10-APR-2001
DEFINITION   602586054P1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4714909 5',
             mRNA sequence.
ACCESSION    BG564257
VERSION      BG564257.1    GI:13571909
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 810)
  AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: CLONETECH Laboratories, Inc.
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM1560  row: 1  column: 14
             High quality sequence stop: 772.
FEATURES             Location/Qualifiers
     source          1..810
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4714909"
                     /clone_lib="NIH_MGC_76"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1:
                     sfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGCCCGAGGCCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C, G, or T). Average insert size 1.85
                     kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

Query Match       27.6%;   Score 611;  DB 12;  Length 810;
Best Local Similarity  92.9%;  Pred. No. 2e-79;
Matches     685;  Conservative     0;  Mismatches     45;  Indels     7;  Gaps     4;
```

```
Qy     61  ATAACTTTTATTCTACTATGTATATGTATGTATGATTATTAATAAATGAACTAGGAAG  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     71  ATAACTTTTTATTCTACTATGTATATGTATGTATGATTATTAATAAATGAACTAGGAAG  130

Qy    121  GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAAGAGATGGAAACAACATTGCTTTTC  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    131  GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAAGAGATGGAAACAACATTGCTTTTC  190

Qy    181  TTTTCTCAAATAAATATGTGAATCAAAAGAAAAAAACTTTTTCAAGTAATACATGGT  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    191  TTTTCTCAAATAAATATGTGAATCAAAAGAAAAAAACTTTTTCAAGTAATACATGGT  250

Qy    241  TCAGGAA-AAGAAGAAACAAGCAAAGAAGCCAAAATACAGAGCTAAGGAAAAAAGAAATAG  299
           ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251  TCAGGAACAAGAAGAAACAAGCAAAGAAGCCAAAATACAGAGCTAAGGAACCAAGACATAG  310

Qy    300  ACTAAGTCTTCTTGTGCAGAAACCAAGAGTCTCCCTGAAGAGAGGCAGTGAAATGGGGTGAATCATT  359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311  ACTAAGTCTTCTTGTGCAGAAACAAGAGTCTCCCTGAAGAGAGGCAGTGAAATGGGGTGAATCTGG  370

Qy    360  GCACTTGGCCAAAGAAACAAGAGTCTCCCTGAAGAGAGGCAGTGAAATGGGGTGAATCATT  419
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    371  GCACTTGGCCAAAGAAACAAGAGTCTCCCTGAAGAGAGGCAGTGAAATGGGGTGAATCATT  430
```

JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
             Group Phase I & II Team.
TITLE   Analysis of the mouse transcriptome based on functional annotation
             of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 1860)
AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
             Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
             Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
             Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
             Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
             Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
             Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
             Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
             Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
             Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
             Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
             Muramatsu,M. and Hayashizaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
             RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
             Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
             Please visit our web site for further details.
             URL:http://genome.gsc.riken.go.jp/
             URL:http://fantom.gsc.riken.go.jp/.
FEATURES   Location/Qualifiers
     source   1..1860
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM DB:9830006J10"
                     /db_xref="MGI:2400834"
                     /db_xref="taxon:10090"
                     /clone="9830006J10"
                     /sex="male"
                     /tissue_type="bone"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS   182..889
                     /note="unnamed protein product; putative
                     regulator of G-protein signaling 18 (MGD|MGI:1927498,
                     GB|NM_022881, evidence: BLASTN, 99%, match=1393)"
                     /codon_start=1
                     /protein_id="BAC29417.1"
                     /db_xref="GI:26313374"
                     /translation="MDMSLVFFSQLNMCESKEKTPFKLMHGSGKEETSIEAKIRAKEK
                     RNRLSLLQRPDFHGETQASRSALLAKETRVSPEAVKWAESPDKLLSHRDGVDAFTR
                     FLKTEFSEENIEFWVACEDFKKCKEPQQILKAKAIYEKFIQNDAPKEVNIDFHTKEV
                     IAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLRHLIEGRPQRPTNLRRRSRS
                     FTYNDFQDVKSDVAIWL"
     polyA_signal   1839..1844
                     /note="putative"
     polyA_site   1860
                     /note="putative"
ORIGIN

Query Match   26.9%;   Score 596;   DB 11;   Length 1860;
Best Local Similarity   72.1%;   Pred. No. 1.8e-77;
Matches 833;   Conservative   0;   Mismatches 290;   Indels 33;   Gaps 3;

Db   89 TTTATTCAGGATGGTGTTTTCCTATGAAATAGCATTCATCTGTGGAGGAGAAGGACTAAGG 148
Qy   69 TTATTCCTACTATGTATATGTGGAATAGTATTAATAAATGAACTAGGGAAGGATGGTAAT 128
          || || || |||| |||||| || ||| || | | | |||

Db   129 AAATTAGACATCTTCAT--TTTAGAGAAGATGGAACAACATTGCTTTCTTTTCT 186

Qy   149 AAATCTGACATCTGTTGGTCACTGGGACAGAATATGTCACTGGTTTTCTCTCT 208

Db   187 CAAATAAATATGTGTGAATCAAAAGAAAAAACTTTTTTCAAGTTAATAACATGGTTCAGGA 246

Qy   209 CAATTAAATATGTGTGAATCAAAAGAAAAAACTTTTTTCAAACTAATCATGGGTCAGGG 268

Db   247 AAAGAGAAACAAGCAAGAGAGCCAAAATCAGAGCTAAGGAAAAAAGAAATAGACTAAGT 306

Qy   269 AAAGAAGAAACAAGCATCGAGGCCAAAATCAGAGCGGCCAAAATCAGAGCGGCCAAAATCAGAGCTAAGT 328

Qy   307 CTTTCTGTGCAGAAACCTGAGTTTCATGAAGACACACCGCTCCAGTAGATCTGGGCACTTG 366

Db   329 CTTTCCCTACGAGAGGCCTGACTTCCATGGAGAGACTCAAGCCCAGTAGAGTCTGCCCTCTTG 388

Qy   367 GCCAAAGAAACAAGAGCTCTCCCCTGAAGAGCAGTGAAATGGGGTGAATCATTTGACAAA 426

Db   389 GCCAAAGAAACAAGAGTGTCTCCTGAAGAGCAGTGAAATGGGCTGAATCATTTGACAAA 448

Qy   427 CTGCTTTCCCATAGAGATGGACTAGAGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 486

Db   449 TTGCTCTCATAGAGATGGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 508

Qy   487 GAAGAAAAATATTGAATTTGGATAGCCTGTGAAGATTCAAGAAAGAAAGCAAGGGACCTCAA 546

Db   509 GAGGAGAACAGTTGAATTTGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAA 568

Qy   547 CAAATTCACCTTAAAGCAAAGCAATATGAGAAATTTATACGACTGATGCCCCAAAA 606

Db   569 CAAATCATCCTAAAAGCAAAGGCAATCTATGAGAAATTCATTCAGAGAAATTCAGAGAATGATGCCCCCAAA 628

Qy   607 GAGGTAACCTGATTTTCACACAAAAGAAGTCATTACAAACACAGCATCCACCTACC 666

Db   629 GAGGTAACATTGATTTTCATCATCAAGAAGTAATTGCTAAGAGCATCGCCCAGCCACT 688

Qy   667 CTCCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAGACAGTTAT 726

Db   689 CTCCACAGTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTAT 748

Qy   727 ACACGTTTCTGAAATCTGACATCTATTAGACTTGATGGAAGAAGACCTCAGAGACCA 786

Db   749 AAACGTTTTTGAAATCTGAAATCTGACCTTACATTTGATGAGAAGGAAGACCTCAGAGACCA 808

Qy   787 ACAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATTACAAATCA 846

Db   809 ACAAACCCTTAGGAGACGATCACGATCATTTACCTTACAATGAATGATTTCCAAGATTGTAAAGTCA 868

Qy   847 GATGTTGCCATTGGTTATAAAGAAAATTGATTTTGCTCATTTTTATGACAAACTTATAC 906

Db   869 GATGTTGCCATTGGTTATGGTAATGAGTAAAAGTCATTTGCTTCTTTTTGTAGTGTATGTGTAT 928

Qy   907 ATC----------------------------TGCTTCTTAACATATCGCATGTTTATGTTAAGA 941

Db   929 ATCTAAAATATATAACTAATACTAATAGTGTACTTCTAAAATAGCTTGTTGTATAAGAAGA 988

Db   942 TTTGGTCCCATCCTTTAAACTGAAATATGTCATGGAGAATTATTTTAAAATGTAAAAAC 1001

Db   989 GATGATTTCATTTT----TAAAATACACCATGCAAATAACATATTACATTGTAAGAAACTT 1042

Qy 1002 AAAACTTTCTGCTAACAAAATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTA 1061

Db 1043 TTTATTATTACTACAAAATAAAATATTCATCATCCTTCCGAAATATTTATGAAAATATCATC 1102

Qy 1062 TTTGATGTCATTCATTCATTTATAATCAGAAAAAAAAACTTATTTCTTAATCAAAAGGCAGTAC 1121

Db 1103 TGATATTCATTCAATAAAATCTTTATTTCTCACAATAACAGTCAGTAAGAAGAGAAGCTT 1162

Qy 1122 AAAAAAGTAATAATGTTTTATAAGACTTGTAGAGTTAAGTAAAGTTAAGCTTTTGCAA 1181

Db 1163 TGTAAATTGTTAAAGTAAAATGAAAAACTTTGTAGTAATGTAAAAATGAACTAGGGAGGGATGGTAAT 1222

Qy   1182 GTTGTCAAAAGTTCAA 1197
          ||| | | ||| |||
Db   1223 TTTCTAAGCAAAGCAA 1238

RESULT 10
AV714060
LOCUS   AV714060   705 bp   mRNA   linear   EST 11-OCT-2000
DEFINITION AV714060 DCB Homo sapiens cDNA clone DCBBIF02 5', mRNA sequence.
ACCESSION   AV714060
VERSION   AV714060.1 GI:10795577
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 705)
AUTHORS   Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Cheng,Z. and Han,Z.
TITLE   Homo sapiens cDNA DCB clones
JOURNAL   Unpublished (2000)
COMMENT   Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES   Location/Qualifiers
     source   1..705
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DCBBIF02"
                     /cell_type="dendritic cells"
                     /dev_stage="mature"
                     /lab_host="BM25.8"
                     /clone_lib="DCB"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
ORIGIN

Query Match   25.6%;   Score 568; DB 9; Length 705;
Best Local Similarity   96.5%;   Pred. No. 3.8e-73;
Matches 580; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy   61 ATAACTTTTATTCTACTATGTATATGTATATGGAATAATAGTATTAAATAATGAACTAGGGAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 ATAACTTTTATTCTACTATGTATATGTATATGGAATAATAGTATTAAATAATGAACTAGGGAGG 130

Qy   121 GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC 180
         |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   131 GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC 190

Qy   181 TTTTCTCAAATAAATAATGTGTGAATCAAAGAAAAGAAAAAACTTTTTCAAGTTAATACATGGT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   191 TTTTCTCAAATAAATAATGTGTGAATCAAAGAAAAGAAAAAACTTTTTCAAGTTAATACATGGT 250

Qy   241 TCAGGAAAAGAAGAAGAAGAAGCAAGAGAGCCAAAATCAGAGCTAAGGAAGAAAAAGAAATAGA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   251 TCAGGAAAAGAAGAAGAAGAAGCAAGAGAGCCAAAATCAGAGCTAAGGAAGAAAAAGAAATAGA 310

Qy   301 CTAAGTCTTCTTGTGCAGAAACCTGAGTTCATGAGAACCACCCGCTTCAGTAGATCTGGG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   311 CTAAGTCTTCTTGTGCAGAAACCTGAGTTCATGAGAACCACCCGCTTCAGTAGATCTGGG 370

Qy   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   371 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTT 430

Qy   421 GACAAACTGCTTCCATAGAGATGGACTAGAGGCTTTACCAGATTTCTTAAAACTGAA 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   431 GACAAACTGCTTCCATAGAGATGGACTAGAGGCTTTACCAGATTTCTTAAAACTGAA 490

Qy   481 TTCAGTGAAGAAAATATTGAATTTGAATTTGATAGCCTGTGAAGATTCAAGAAAAGCAAGGGA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   491 TTCAGTGAAGAAAATATTGAATTTGAATTTGATAGCCTGTGAAGATTCAAGAAAAGCAAGGGA 550

Qy   541 CCTCAACAAATCACCTTAAAGCAAAGCAATATATGAGAAATTATACAGACTGATGCC 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   551 CCTCAACAAATCACCTTAAAGCAAAGCAATATATGAGAAATTATACAGACTGATGCC 610

Qy   601 CCAAAAGAGGTTAACCTTGATTTTCCACAAAAGAGTCATTACAAACAGCAGACTCACTCAA 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   611 CCAAAAGAGGTTAACCTTGATTTTCCACACTGCACAAACCACANAGCAGAGTGTATCAGCTCATTGAA 670

Qy   661 C 661
         |
Db   671 C 671

RESULT 11
AV716055
LOCUS   AV716055   639 bp   mRNA   linear   EST 11-OCT-2000
DEFINITION AV716055 DCB Homo sapiens cDNA clone DCBBOD01 5', mRNA sequence.
ACCESSION   AV716055
VERSION   AV716055.1 GI:10797572
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 639)
AUTHORS   Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Cheng,Z. and Han,Z.
TITLE   Homo sapiens cDNA DCB clones
JOURNAL   Unpublished (2000)
COMMENT   Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES   Location/Qualifiers
     source   1..639
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DCBBOD01"
                     /cell_type="dendritic cells"
                     /dev_stage="mature"
                     /lab_host="BM25.8"
                     /clone_lib="DCB"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
ORIGIN

Query Match   25.5%;   Score 565.4; DB 9; Length 639;
Best Local Similarity   99.5%;   Pred. No. 9.6e-73;
Matches 566; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   61 ATAACTTTTATTCTACTATGTATATGTATATGGAATAATAGTATTAAATAATGAACTAGGGAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 ATAACTTTTATTCTACTATGTATATGTATATGGAATAATAGTATTAAATAATGAACTAGGGAAG 130

Qy   121 GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 GATGTAATAAATTAGACATCTCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC 190

Qy   181 TTTTCTCAAATAAATAATGTGTGAATCAAAGAAAAGAAAAAACTTTTTCAAGTTAATACATGGT 240

Matches 566; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy    61 ATAACTTTTATTCTACTATGTATATGTATGTGGAATAGTATTAATAAATGAACTAGGGAAG 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    58 ATAACTTTTATTCTACTATGTATATGTATGTGGAATAGTATTAATAAATGAACTAGGGAAG 117

Qy   121 GATGTAATAAATTAGACATCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   118 GATGTAATAAATTAGACATCTTCATTTTAGAGAGAAGATGGAAACAACATTGCTTTTC 177

Qy   181 TTTTCTCAAATAAATATGTGAATCAAAAGAAAAAACTTTTTCAAGTTAATACATGGT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   178 TTTTCTCAAATAAATATGTGAATCAAAAGAAAAAACTTTTTCAAGTTAATACATGGT 237

Qy   241 TCAGGAAAAGAAGAAACAAGAAGCCAAAATCAGAGCTAGGGAAAAAGAAAATAGA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   238 TCAGGAAAAGAAGAAACAAGAAGCCAAAATCAGAGCTAGGGAAAAAGAAAATAGA 297

Qy   301 CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGG 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   298 CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGG 357

Qy   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTT 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   358 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAAATGGGTGAATCATTT 417

Qy   421 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   418 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA 477

Qy   481 TTCAGTGAAGAAATATTGGATAGCCTGTGAAGATTTCGAGATTAGCCTGTGAAGATTTCGAGATTAGCCTGTGAAGATAAGCAAGGGA 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   478 TTCAGTGAAGAAATATTGGATAGCCTGTGAAGATTTCGAGATTAGCCTGTGAAGATAAGCAAGGGA 537

Qy   541 CCTCAACAAATTCACCTTAAAGCAAAAGCAATAATATGAGAAATTTATACAGACTGATGCC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   538 CCTCAACAAATTCACCTTAAAGCAAAAGCAATAATATGAGAAATTTATACAGACTGATGCC 597

Qy   601 CCAAAAGAGAGTTAACCTTGATTTTCACAC 629
            ||||||||||||||||||||||||||||||
Db   598 CCAAAAGAGAGTTAACCTTGATTTGTCACAC 626
```

RESULT 13
CF181629
LOCUS   CF181629   818478 MARC 3PIG Sus scrofa cDNA 5', mRNA sequence.   723 bp   mRNA   linear   EST 28-JUL-2003
DEFINITION 818478 MARC 3PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   CF181629
VERSION   CF181629.1 GI:33293405
KEYWORDS   EST.
SOURCE   Sus scrofa (pig)
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1 (bases 1 to 723)
AUTHORS   Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Fox,J.,
            Wise,T.A., Nonneman,D.J., Wray,J.E. and Keele,J.W.
TITLE   A second set of porcine ESTs from a pooled-tissue normalized
            library
JOURNAL   Unpublished (2003)
COMMENT   Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: SRG8012 row: O column: 18
            seq primer: GTAATACGACTCACTATAGGG.
FEATURES   Location/Qualifiers
     source   1..723

```
Db   191 TTTTCTCAAATAAATATGTGAATCAAAAGAAAAAACTTTTTCAAGTTAATACATGGT 250
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   241 TCAGGAAAGAAGAAACAAGAGCCAAATCAGAGCTAGGCAAATCAGAGCTAAGGAAAAAGAAAATAGA 300

Qy   251 TCAGGAAAGAAGAAACAAGCAAGAGCTAAGGAAAAAGAAAAAGAAAATAGA 310

Db   301 CTAAGTCTTCTTGTGCAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   311 CTAAGTCTTCTTGGCAGAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGG 370

Db   361 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAATGGGTGAATCATTT 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   371 CACTTGGCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCAGTGAATGGGTGAATCATTT 430

Db   421 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   431 GACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAA 490

Db   481 TTCAGTGAAGAAATATTTGGATATTTGGATAGCAATAATAGAGAAATTTATACAGACAAGGGA 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   491 TTCAGNGAAGAAATATTTGGATAGCAATAATAGAGAAATTTATACAGACAAGGGA 550

Db   541 CCTCAACAAATTCACCTTAAAGCAAAAGCAATAATATGAGAAATTTATACAGACTGATGCC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   551 CCTCAACAAATTCACCTTAAAGCAAAGCAATAATATGAGAAATTTATACAGACTGATGCC 610

Db   601 CCAAAAGAGAGTTAACCTTGATTTTCACAC 629
            ||||||||||||||||||||||||||||||
Qy   611 CCAAAAGAGAGTTAACCTTGATTTTCACAC 639
```

RESULT 12
AV714044
LOCUS   AV714044   626 bp   mRNA   linear   EST 11-OCT-2000
DEFINITION AV714044 DCB Homo sapiens cDNA clone DCBCFH06 5', mRNA sequence.
ACCESSION   AV714044
VERSION   AV714044.1 GI:10795561
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 626)
AUTHORS   Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Gu,Y., Yang,Y., Gao,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Cheng,Z. and Han,Z.
TITLE   Homo sapiens cDNA DCB clones
JOURNAL   Unpublished (2000)
COMMENT   Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES   Location/Qualifiers
     source   1..626
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DCBCFH06"
                     /cell_type="dendritic cells"
                     /dev_stage="mature"
                     /lab_host="BM25.8"
                     /clone_lib="DCB"
                     /note="Vector: pTrip1Ex2; Site_1: sfiIA; Site_2: sfiIB"
ORIGIN

Query Match   25.4%;   Score 564.2;   DB 9;   Length 626;
Best Local Similarity   99.5%;   Pred. No. 1.4e-72;

```
                /organism="Sus scrofa"
                /mol_type="mRNA"
                /db_xref="taxon:9823"
                /tissue_type="pooled"
                /lab_host="DH10B"
                /clone_lib="MARC 3PIG"
                /note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
                Library made with RNA pooled from multiple tissues
                including brain, liver, muscle, placenta/endometrium,
                ovary, testes, and bone marrow."

ORIGIN

Query Match      25.2%;  Score 557.6; DB 14;  Length 723;
Best Local Similarity  87.3%;  Pred. No. 1.2e-71;
Matches    623;  Conservative    0;  Mismatches    89;  Indels    2;  Gaps    1;

Qy        107 ATGAACTAGGGAAGGATGTAATAAAATTAGACATCTCTTCA--TTTTAGAGAGAAGATGGA 164
              |||||| |||||||| ||| ||| |||||||||||||||||  |||||||||||||||||
Db          9 ATTAACAAGTGGGAAGATGTAATAAATTATACAACTCTTCACTTTATAGAGAGAAGATGGA 68

Qy        165 AACAACATTGCTTTCTTTTCTCAAATAAATATGTGAATCAAAGAAAAAACTTTTTT 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 TACACCCCTGCTTTTCTTTTTTTCAATTAAATATGTGCGAACCAAAGAAAAAACTTTTTT 128

Qy        225 CAAGTAATACATGGTTCAGGAAAAGAAGAAGAAACAAGCAAAGAAGCAAAATCAGAGCTAA 284
              |||| |||||||||||||||| |||||||||||||||||||||||||| ||||||||||||
Db        129 CGAGTTAATTCATGGTTCAGGAAAAGAAGAAGAAACAAGCAAAGAAGCCAGAATCAGAGATAA 188

Qy        285 GGAAAAAGAAATAGACTAAGTCTTCTTGTGTGCAGAAACCTGAGTTTCATGAAGACACCCG 344
              |||||||||||||||||||||  |||| |||||||||||||||||||||||||||||||||
Db        189 GGAAAAAGAAACAGGCTAAGTCTTCTTCCCTGCGAAGAAACAAGAGTCTCCCTGAAGAGCAGTGAA 248

Qy        345 CTCCGGTAGATCGGCCACTTGGGCACTTGGCCAAAGAAACAAGAGTCTCCCTGAAGAGGCAGTGAA 404
              |||||| ||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        249 CTCCGGTAGTTGTTCTGAACACCTGCGCAAAAAAACAAGAATCTCCCCGAAGAAGCAATGAT 308

Qy        405 ATGGGGTGAATCATTGACAAACTGCTTTCCCATAGAGATGGACTAGAGGCTTTTACCAG 464
              |||||| ||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        309 ATGGGGTGAATCATTCATTGACAAACTGCTTTCCCGAAAAGATGGACTTTTCACCGG 368

Qy        465 ATTTCTTAAAACTGAAACTCAGTGAAGAAAATATTGAATTTTGGATAGCCGTGGAAGATTT 524
              |||||||||||||||||| |||||| ||||||| ||||||||||||||||||||||||||||
Db        369 ATTTCTTAAGACCGAATTCAGTGATGATAACATTGCATTTTGGATAACATTGCATTTTGGATAGCCGTGAAGATTA 428

Qy        525 CAAGAAAAGCAAGGGACCTCAACAAATCACCTTAAAGCAAAAGCAATATATGAGAAAATT 584
              |||||||||||||||||||||||||||| |||||||| |||||| |||||||||||||||
Db        429 CAAGAAAAGCAAGGGACCCTCAACAAATAATTCGAAAGCGAAAGCAATATATGAGACGTT 488

Qy        585 TATACAGACTGATGCCCCAAAAGAGGTTAACCTTGATTTTCACACAAAAGAAGTCATTAC 644
              ||||||||||||||  |||||||||||| ||||||||||||||||||| ||||| |||||
Db        489 TATACAGACCGATGCTCCACAAGATGTTAACCTTGATTTTCACACCAAAGAAAATCCTCGC 548

Qy        645 AAACAGCATCACTCAACCTACCCTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCA 704
              ||||||||||| |||||||||||||||||||||||||||||||| ||||| |||||||||
Db        549 CAAGATCATCACGCAGCCTACCCTCCACAGCTTTGATGCTGCCATAGCCATAGTGTATCA 608

Qy        705 GCTCATGGAACAAGCAGCAGTTATACACGTTTCTGAAATCTGACATCTATTAGACTTGAT 764
              |||||||||||||||||||||| ||||||||||||||||||| |||||||||||||||||
Db        609 TCTCATGGAACAAGCAGCAGTTACACGTTTTCTGAAATCCGACATCTATTAGACTTGAT 668

Qy        765 GGAAGGAGACCTCAGAGACCAACAACAAATCTTAGGAGACGATCACGCTCATTAC 818
              |||||| ||||||||||||||||||||||||||||||||| |||||||||||||||
Db        669 AGAAAGAGACCTCGGAGCGACGAACAAATCTGAGAGGCGATCACGTTCCTTTAC 722


RESULT 14
CD466872
LOCUS       CD466872     LeukoN2_8_B10_g1_A024  661 bp    mRNA     linear    EST 04-JUN-2003
DEFINITION  LeukoN2_8_B10_g1_A024 Unstimulated peripheral blood leukocytes N2
            Equus caballus cDNA clone LeukoN2_8_B10_A024 5', mRNA sequence.
ACCESSION   CD466872
VERSION     CD466872.1  GI:31388140
KEYWORDS    EST.
```

SOURCE   Equus caballus (horse)
  ORGANISM Equus caballus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE   1 (bases 1 to 661)
  AUTHORS   Vandenplas,M., Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S.,
            Moore,J.N., Liang,C., Sun,F., Sullivan,R., Shah,M. and Pratt,L.H.
  TITLE   An EST database from equine (Equus caballus) unstimulated
            peripheral blood leukocytes
  JOURNAL   Unpublished (2003)
COMMENT   Other_ESTs: LeukoN2_8_B10.b2_A024
            Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
            the Human Genome Center, University of Tokyo Institute of Medical
            Science; tissue and RNA were prepared in the Department of Large
            Animal Medicine, University of Georgia; sequencing done in the
            Laboratory for Genomics and Bioinformatics, University of Georgia.
            Sequence ends have been trimmed to exclude vector and regions below
            Phred quality 16. Three-prime sequences are presented as their
            reverse complement and have been trimmed to exclude polyA.
            Seq primer: Sug5 (CTTCTGCTCTAAAAGCTGCG).
FEATURES   Location/Qualifiers
     source   1..661
                     /organism="Equus caballus"
                     /mol_type="mRNA"
                     /strain="thoroughbred"
                     /db_xref="taxon:9796"
                     /clone="LeukoN2_8_B10_A024"
                     /sex="male"
                     /tissue_type="blood"
                     /cell_type="leukocytes"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="Unstimulated peripheral blood leukocytes N2"
                     /note="Organ: circulatory system; Vector: pME18S-FL3;
                     Site 1: XhoI; Site 2: XhoI; The library was prepared from
                     polyA+ RNA from unstimulated equine peripheral blood
                     leukocytes isolated from a healthy adult horse.
                     Double-stranded cDNA was cloned unidirectionally into
                     different DraIII sites of the pME18S-FL3 vector (5-prime
                     DraIII site is CACTGTGTG, 3-prime DraIII site is
                     CACCATGTG). XhoI excises the cDNA insert."

ORIGIN

Query Match   24.8%; Score 550; DB 14; Length 661;
Best Local Similarity 90.6%; Pred. No. 1.6e-70;
Matches   598; Conservative   0; Mismatches   60; Indels   2; Gaps   1;

```
Qy        121 GATGTAATAAATTAGACATCTCTTCATTTTAGAG--AGAAGATGGAAGAACAACATTGCTTT 178
              |||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||||
Db          2 GGTGTAATAAATTAGAGTGTCTCTTCATTTTAGAGAGAGAAGAAGATGGAAACATCATTGGTTT 61

Qy        179 TCTTTTCTCAAATAAATATGTGAATCAAAGAAAAAACTTTTTCAAGTAATAACATG 238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TCTTTTCTCAATTAAATATGTGAATCAAAGAAAAAACTTTTTTCAAGTAATAACATG 121

Qy        239 GTTCAGGAAAAGAAGAAGAAACAAGCAAAGAAGCAAAATCAGAGCTAAGGAAAAAGAAATA 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GTTCAGGAAAAGAAGAAGAAACAAGCAAAGAAGCAAAACTCAGAGCTAAGGAAAAAGAAACA 181

Qy        299 GACTAAGTCTTCTTGTGTGCAGAAACCTGAGTTTCATGAAGACACACCCGTCCAGTAGATCTG 358
              ||||||||||||||||||||||||||||||||||||||||||||||  |||||||||||||
Db        182 GGTTAAGTCTGCTCTTGCAGAAACCTGAATTTCATGAAGAGACCCCACTCCGGTAGATCCG 241

Qy        359 GGCACTGGCCAAGAAACAAGAGTCTCCCTGAAGAGGCAGTGAAATGGGGTGAATCATG 418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GGCACTTAGCCAAGAAACGAGAATCTCCCCGTGCAGAAGCCGTGAATATGGGGTGAATCAT 301
```

```
Best Local Similarity  100.0%;  Pred. No. 1.4e-68;
Matches   537;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   829 TTCCAAGATGTACAATCAGATGTTGCCATTGGTATAAAGAAAATTGATTTGCTCATT  888
         ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TTCCAGAGTGTACAATCAGATGTTGCCATTGGTATAAAGAAAATTGATTTGCTCATT   60

Qy   889 TTTATGACAAACTTATACATCGTCTCTAACATATCGCATGTTATGTTAAGATTTGGTC  948
         |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    61 TTTATGACAAACTTATACATCGTCTCTAACATATCGCATGTTTATGTTAAGATTTGGTC  120

Qy   949 CCATCCTTAAACTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAACAAAACTT 1008
         ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   121 CCATCCTTAAACTGAAATATGTCATGTGAAATTA TTTTAAAAATGTAAAAACAAAACTT  180

Qy  1009 TCTGCTAACAAAATACATACAGTATCTGCCAGTATATTCGTAAAACCTTCTATTTGATG 1068
         ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TCTGCTAACAAAATACATACAGTATCTGCCAGTATATTCGTAAAACCTTCTATTTGATG  240

Qy  1069 TCATTCCATTTATATCAGAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAA 1128
         ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   241 TCATTCCATTTATATCAGAAAAAAAACTTATTTCTTAATCAAAAGGCAGTACAAAAAAA  300

Qy  1129 GTAATAATGTTTTTATAAGATGTAGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTGTCA 1188
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GTAATAATGTTTTTATAAGATGTAGAGTTAAGTAAAAGTTAAGCTTTGCAAAGTGTCA  360

Qy  1189 AAAGTTCAAACAAAAGTCTAGTTGGGATTTTTACCAAGCAGCATATATGATGTTATAT 1248
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AAAGTTCAAACAAAAGTCTAGTTGGGATTTTTACCAAGCAGCATATATGATGTTATAT  420

Qy  1249 AAACATAATAATACTTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCT 1308
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AAACATAATAATACTTCAGATATCCAAATGTTCAGATAGCATTTTCATAATGAATGTTCT  480

Qy  1309 CTTTTTTTGGTAATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAATGAAGAAC 1365
         |||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   481 CTTTTTTTGGTAATAGTGTAGAAGTGATCTGGTTCTTACAATGGGAGATGAAGAAC  537
```

Search completed: August 20, 2004, 12:31:50
Job time : 6017.98 secs

```
Qy   419 TTGACAAACTGCTTTCCATAGAGATGGACTAGAGGCTTTACCAGATTTCTTAAAACTG  478
         |||||||||||| | ||||| || ||||||  ||| ||||||||||||||||||||| |
Db   302 TTGACAAACTGCTCCCCATAAAGGTGGACTGGGACTTTTACCAGATTTCTTAAAACCG  361

Qy   479 AATTCAGTGAAGAGAAAATATTGAATTTGATTT...                        538
         |||||||||| | | |||||| |  ||||||||
Db   362 AATTCAGTGACAATCAGAGTGCCATTGGTATAAAGAAAATGATTTGCTCATT        421

Qy   539 GACCTCAACAAATTCACCTTAAAGCAAAGCAATATATGACAAAGAATTATACAGACTGATG  598
         ||| ||||||||  | ||||||||||||| ||||||||| |||| ||||||||||  ||||
Db   422 ACCCTCAACAAAATATCCTTAAAGCAAAGTAATATATGAGAAATTATACAGAACGATG     481

Qy   599 CCCCAAAAGAGAGTTAACCTTGATTTCACACAAAGAAGTCATTACAAACAGCATCACTC  658
         | || |||||||||||||||||||||| |||| |||||| ||||| | | ||||||| |
Db   482 CTCCGCAAGAGAGTTAACCTTGATTTCCACACCAAAGAAATCATTGCCAAGAGCATCACCA  541

Qy   659 AACCTACCCTCCACAGTTTGATGCTGCACAAAGCACGAGTGTATCAGCTCATGGACAAG  718
         |||| || ||||||||||||||||||| |||||| |||| |||||||| |||||||||
Db   542 AACCCACTCTCCACAGTTTGATGCTGCCACAAAGCAGAGTATATCAGCTTATGGACAAG  601

Qy   719 ACAGTTATACAGTTTCTGAAATCGACATCTATTAGACTGATGGAAGGACCTC  778
         |||||||  |||||||||||||| |||||||||||||||| || |||||||||
Db   602 ACAGTTACACAGTTTCTGAAATCTGACATCTATTAGACTGACAGAGAAGGAAGACCTC  661
```

RESULT 15
AQ317367
LOCUS   AQ317367   537 bp   DNA   linear   GSS 04-MAY-1999
DEFINITION RPCI11-104A6.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-104A6,
            genomic survey sequence.
ACCESSION   AQ317367
VERSION   AQ317367.1 GI:4048618
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 537)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE   Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL   Unpublished (1998)
COMMENT   Other GSSs: RPCI11-104A6.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES   Location/Qualifiers
     source   1..537
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7539557"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-104A6"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN

Query Match   24.2%; Score 537; DB 28; Length 537;

```
                GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      August 20, 2004, 12:36:30 ; Search time 79 Seconds
                                 (without alignments)
                          840.489 Million cell updates/sec

Title:           US-09-894-749-2
Perfect score:   1218
Sequence:        1 METTLLFFSQIMCESXEKT..........SRSFTCNEFQDVQSDVAIWL 235

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:    1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_29Jan04:*
                 1: geneseqp1980s:*
                 2: geneseqp1990s:*
                 3: geneseqp2000s:*
                 4: geneseqp2001s:*
                 5: geneseqp2002s:*
                 6: geneseqp2000as:*
                 7: geneseqp2003bs:*
                 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                                SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1218	100.0	235	3	AAY97153	Aay97153 Human reg
2	1218	100.0	235	5	AAE25829	Aae25829 Human reg
3	1218	100.0	235	5	AAU10749	Aaul0749 Human RGS
4	1218	100.0	235	5	ABP64801	Abp64801 Human pro
5	1176	96.6	227	5	ABP64054	Abp64054 Human ORF
6	1032	84.7	235	3	AAY97154	Aay97154 Murine re
7	1032	84.7	235	5	AAE25830	Aae25830 Murine re
8	422.5	34.7	211	6	ABP96779	Abp96779 Human COP
9	422.5	34.7	211	6	ABP98020	Abp98020 Amino aci
10	422.5	34.7	211	7	ADD14019	Add14019 Human src
11	422.5	34.7	220	3	AAB56929	Aab56929 Human pro
12	421.5	34.6	181	2	AAW62075	Aaw62075 Human reg
13	421.5	34.6	181	3	AAY53931	Aay53931 A human r
14	421.5	34.6	190	5	ABP62144	Abp62144 Human sec
15	413	33.9	81	5	AAU10747	Aaul0747 Amino aci
16	410.5	33.7	181	5	ABB83788	Abb83788 Rat disea
17	409	33.6	930	6	ABB99465	Abb99465 Amino aci
18	408.5	33.5	205	5	AAU78976	Aau78976 Human RGS
19	408.5	33.5	205	7	ADE62388	Ade62388 Human Pro
20	408.5	33.5	205	5	ADE62392	Ade62392 Human Pro
21	405	33.3	198	4	AAB62187	Aab62187 Human RGS
22	404.5	33.2	205	7	ADE62386	Ade62386 Rat Prote
23	404.5	33.2	205	7	ADE62390	Ade62390 Rat Prote
24	404	33.2	519	4	AAB90786	Aab90786 Human she
25	404	33.2	519	6	ABB99466	Abb99466 Amino aci
26	404	33.2	917	6	ADA13333	Ada13333 Human int
27	402.5	33.0	180	5	ABP69254	Abp69254 Human pol
28	402.5	33.0	180	5	ABB04999	Abb04999 Human reg
29	400	32.8	284	6	ADA54082	Ada54082 Human pro
30	396.5	32.6	923	6	ABU54617	Abu54617 Human NOV
31	381	31.3	135	6	ABP98021	Abp98021 Amino aci
32	381	31.3	152	4	AAG66509	Aag66509 RGS prote
33	380	31.2	202	2	AAR89895	Aar89895 p53 respo
34	380	31.2	202	2	AAW08133	Aaw08133 Human cyt
35	380	31.2	202	2	AAW59294	Aaw59294 Human RAT
36	380	31.2	202	2	AAY87952	Aay87952 Human CR1
37	380	31.2	202	6	ABR92146	Abr92146 Human cer
38	370	30.4	128	7	ABR82620	Abr82620 Rat RGS4
39	370	30.4	167	5	AAU11043	Aaul1043 Rat RGS4
40	369	30.3	201	2	AAW59293	Aaw59293 Mouse RAT
41	367	30.1	118	2	AAY41001	Aay41001 Mouse RGS
42	361	29.6	69	5	AAU10748	Aaul0748 C-termina
43	361	29.6	116	2	AAY41020	Aay41020 RGS2 prot
44	357	29.3	249	6	ABP98019	Abp98019 Amino aci
45	353	29.0	196	4	AAB76863	Aab76863 Human lun

```
                               ALIGNMENTS

RESULT 1
AAY97153
ID   AAY97153 standard; protein; 235 AA.
XX
AC   AAY97153;
XX
DT   04-DEC-2000  (first entry)
XX
DE   Human regulator of G-protein signaling protein.
XX
KW   RGS; regulators of G-protein signaling; GTPase activating protein; GAP;
KW   G-alpha protein; cell adhesion; chemotaxis; vulnerary; immunosuppressor;
KW   anti-rheumatic; anti-arthritic; anti-diabetic; anti-inflamatory;
KW   cytostatic; hepatotropic; anti-anaemic; modulator; gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key                Location/Qualifiers
FT   Domain             82..201
FT                      /label= RGS_domain
FT   Misc-difference    83
FT                      /note= "forms part of hydrophobic core"
FT   Misc-difference    90
FT                      /note= "forms part of hydrophobic core"
FT   Misc-difference    100
FT                      /note= "forms part of hydrophobic core"
FT   Misc-difference    103
FT                      /note= "forms part of hydrophobic core"
FT   Misc-difference    104
FT                      /note= "forms part of hydrophobic core"
FT   Misc-difference    107
FT                      /note= "forms part of hydrophobic core"
FT   Misc-difference    109
FT                      /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference    111
FT                      /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference    112
FT                      /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference    115
FT                      /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference    116
FT                      /note= "forms part of hydrophobic core"
FT   Misc-difference    138
FT                      /note= "forms part of hydrophobic core"
FT   Misc-difference    139
FT                      /note= "forms part of hydrophobic core"
FT   Misc-difference    142
FT                      /note= "forms part of hydrophobic core"
```

```
FT   Misc-difference 143
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 151
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 152
FT                   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 154
FT                   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 183
FT                   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 184
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 187
FT                   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 188
FT                   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 189
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 191
FT                   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 192
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 193
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 198
FT                   /note= "forms part of hydrophobic core".
XX
PN   WO200046236-A2.
XX
PD   10-AUG-2000.
XX
PF   04-FEB-2000; 2000WO-US002977.
XX
PR   04-FEB-1999;  99US-00244314.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Hodge MR,  Yowe D;
XX
DR   WPI; 2000-532893/48.
DR   N-PSDB; AAA52089.
XX
PT   Novel regulator of G-protein signaling nucleic acids and polypeptides,
PT   useful as diagnostic and investigative tools and to treat G-protein
PT   signaling disorders.
XX
PS   Claim 8; Page 102; 105pp; English.
XX
CC   The RGS (regulators of G-protein signaling) protein genes, clones
CC   AAH16395 and m1975, were identified in human and murine spleen cDNA
CC   libraries, respectively. Both proteins have unique N- and C-terminal
CC   sequences. The C-terminal location of the RGS domain is consistent with
CC   RGSs known to act as GTPase activating proteins (GAPs) for G-alpha
CC   proteins. G-alpha-i linked receptors support rapid adhesion and directed
CC   migration of leukocytes and other cell types. The novel RGS proteins may
CC   be used to modulate cell adhesion and chemotaxis, e.g. for aiding wound
CC   repair. The RGS proteins, related cDNAs and anti-RGS antibodies are
CC   useful for modulation, diagnosis and treatment of immune and respiratory
CC   disorders
XX
SQ   Sequence 235 AA;

Query Match             100.0%;  Score 1218;  DB 3;  Length 235;
Best Local Similarity   100.0%;  Pred. No. 3.5e-111;
Matches  235;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
```

```
Qy  121 DFKKSKGPQQIHLKAKAIYEKPIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 DFKKSKGPQQIHLKAKAIYEKPIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235

RESULT 2
AAE25829
ID   AAE25829 standard; protein; 235 AA.
XX
AC   AAE25829;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Human regulator of G-protein signalling (RGS) protein, h16395.
XX
KW   Human; screening; RGS; regulator of G-protein signalling; pancreatitis;
KW   inflammatory disorder; allergy; Grave's disease; arthritis; sinusitis;
KW   respiratory disorder; asthma; pneumonia; therapy; immune disorder;
KW   haematological disorder; haematopoiesis; platelet-associated disorder;
KW   thrombocytopaenia; migration; invasive disorder; leukaemia; anaemia;
KW   erythrocyte-associated disorder; hepatitis; h16395.
XX
OS   Homo sapiens.
XX
FH                    Location/Qualifiers
FT   Domain           82..201
FT                    /note= "RGS domain"
FT
XX
PN   US6410240-B1.
XX
PD   25-JUN-2002.
XX
PF   04-FEB-2000; 2000US-00498959.
XX
PR   04-FEB-1999;  99US-00244314.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Hodge MR,  Yowe D;
XX
DR   WPI; 2002-588886/63.
DR   N-PSDB; AAD42497.
XX
PT   Screening assays for identifying agent that binds to human or mouse RGS
PT   (regulators of G-protein signalling) protein or its variant, or RGS
PT   protein encoded by homologous DNA sequences, or to host cell expressing
PT   the proteins.
XX
PS   Claim 1; Col 55-56; 42pp; English.
XX
CC   The invention relates to screening assays for identifying an agent that
CC   binds to: human or mouse RGS (regulators of G-protein signalling) protein
CC   or its variant, or RGS protein encoded by homologous DNA sequences; or a
CC   host cell expressing the RGS protein or its variant, or RGS protein
CC   encoded by homologous DNA sequences. The invention is useful for
CC   identifying an agent that binds to human or mouse RGS protein or its
CC   variant, or RGS protein encoded by homologous DNA sequences, or a host
CC   cell expressing the RGS protein or its variant, or RGS protein encoded by
CC   homologous DNA sequences. The agents identified using the invention are
CC   useful for modulating the activity of RGS proteins and thus useful for
CC   treating immune and inflammatory disorders (Grave's disease, allergy,
CC   arthritis), respiratory disorders (asthma, pneumonia, sinusitis),
CC   haematological disorders (haematopoiesis, migration), platelet-associated
CC   disorders (thrombocytopenia), invasive disorders (leukaemia),
CC   erythrocyte-associated disorders (anaemia), pancreatitis, hepatitis etc.
CC   The present sequence is human RGS protein, h16395
XX
SQ   Sequence 235 AA;
```

```
Query Match          100.0%;     Score 1218;  DB 5;   Length 235;
Best Local Similarity  100.0%;   Pred. No. 3.5e-111;
Matches 235;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED  60

Qy  61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy 121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy 181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
```

RESULT 3
AAU10749
ID   AAU10749 standard; protein; 235 AA.
XX
AC   AAU10749;
XX
DT   12-MAR-2002 (first entry)
XX
DE   Human RGS18 polypeptide.
XX
KW   Human; regulator of G protein signalling; RGS18; arterial thrombosis;
KW   platelet activation dysfunction; myocardial infarction; stroke;
KW   coronary artery disease; cerebrovascular disease; unstable angina;
KW   deep vein thrombosis; systemic thromboembolism; anti-coagulant;
KW   invasive cardiac procedure.
XX
OS   Homo sapiens.
XX
XX
PN   WO200183514-A2.
XX
PD   08-NOV-2001.
XX
PF   26-APR-2001; 2001WO-US013540.
XX
PR   28-APR-2000; 2000US-0200786P.
PR   02-AUG-2000; 2000GB-00018833.
XX
PA   (AVET) AVENTIS PHARM PROD INC.
XX
PI   Murray DL, Gagnon AW;
XX
DR   WPI; 2002-055453/07.
DR   N-PSDB; AAS18340.
XX
PT   Isolated regulator of G protein signaling polypeptide, useful for
PT   prevention/treatment of platelet activation dysfunction, such as arterial
PT   thrombosis, myocardial infarction, coronary artery disease and stroke.
XX
PS   Claim 31; Fig 1; 127pp; English.
XX
CC   The present invention relates to the isolation of polynucleotide
CC   sequences, that encode a novel regulator of G protein signalling (RGS)
CC   polypeptide, RGS18, from human platelets. The invention also provides
CC   nucleotide primers and probes specific for an RGS18 nucleic acid. The
CC   sequences of the invention are useful for the manufacture of a medicament
CC   for the prevention or treatment of a platelet activation dysfunction,
CC   such as arterial thrombosis, myocardial infarction, coronary artery
CC   disease, stroke, cerebrovascular disease, unstable angina, deep vein
CC   thrombosis, systemic thromboembolism, and in invasive cardiac procedures
CC   for anti-coagulant purposes The sequences of the invention can also be
CC   used for the detection of RGS18 nucleic acids and RGS18 polypeptides. The
CC   present sequence represents the RGS18 polypeptide isolated from human
CC   platelets
XX
SQ   Sequence 235 AA;

```
Query Match          100.0%;     Score 1218;  DB 5;   Length 235;
Best Local Similarity  100.0%;   Pred. No. 3.5e-111;
Matches 235;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED  60

Qy  61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy 121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy 181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
```

RESULT 4
ABP64801
ID   ABP64801 standard; protein; 235 AA.
XX
AC   ABP64801;
XX
DT   25-FEB-2003 (first entry)
XX
DE   Human protein SEQ ID 461.
XX
KW   Human; expressed sequence tag; EST; haematopoietic disorder;
KW   central nervous system disease; viral infection;
KW   peripheral nervous system disease; non-healing wound; infectious disease;
KW   immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW   fungal infection; autoimmune disorder; coagulation disorder; nootropic;
KW   antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW   cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW   immunostimulant; cerebroprotective.
XX
OS   Homo sapiens.
XX
PN   WO200259260-A2.
XX
PD   01-AUG-2002.
XX
PF   16-NOV-2001; 2001WO-US042950.
XX
PR   17-NOV-2000; 2000US-0071493 6.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI   Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR   WPI; 2002-590824/63.
DR   N-PSDB; ABQ99387.
XX
PT   New isolated polynucleotide, useful in research, diagnostic or
PT   therapeutic methods, e.g. preventing or treating disorders involving
PT   aberrant protein expression or biological activity.
XX
PS   Claim 20; SEQ ID NO 461; 394pp; English.
XX
CC   The present invention relates to novel human coding sequences (ABQ99268-
CC   ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC   therapeutic, diagnostic and research methods. The polynucleotides may be
CC   used in the field of molecular biology as hybridisation probes, primers

CC   for PCR, for chromosome and gene mapping, for the recombinant production
CC   of protein, or in generation of anti-sense DNA or RNA. The
CC   polynucleotides are useful in diagnostics as expressed sequence tags
CC   (ESTs) for identifying expressed genes or for physical mapping of the
CC   human genome. The proteins may be used as molecular weight markers, or as
CC   nutritional sources or supplements. The proteins may be used to maintain
CC   and expand cell population in a totipotential or pluripotential state
CC   useful for re-engineering damaged or diseased tissues, transplantation,
CC   manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC   polynucleotides and proteins are useful for preventing, treating or
CC   ameliorating disorders involving aberrant protein expression or
CC   biological activity. e.g. haematopoietic disorders, central/peripheral
CC   nervous system diseases, mechanical and traumatic disorders, non-healing
CC   wounds, immune deficiencies and disorders, infectious diseases caused by
CC   viral, bacterial or fungal infection, autoimmune disorders, allergic
CC   reactions and conditions, coagulation disorders, or cancer. The
CC   polynucleotide sequences of the invention were assembled from ESTs
CC   isolated mainly by sequencing by hybridisation, and in some cases,
CC   sequences obtained from one or more public databases. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 235 AA;

Query Match   100.0%; Score 1218; DB 5; Length 235;
Best Local Similarity   100.0%; Pred. No. 3.5e-111;
Matches   235; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy   121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy   181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235

RESULT 5
ABP64054
ID   ABP64054 standard; protein; 227 AA.
XX
AC   ABP64054;
XX
DT   04-NOV-2002 (first entry)
XX
DE   Human ORF424.
XX
KW   Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW   Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW   human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW   cancer; cardiovascular disease; allergy; autoimmune disease;
KW   wound healing; blood coagulation disorder; inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   US2002082206-A1.
XX
PD   27-JUN-2002.
XX
PF   30-MAY-2001; 2001US-00867550.
XX
PR   30-MAY-2000; 2000US-0208427P.
XX
PA   (LEAC/) LEACH M D.

PA   (MEHR/) MEHRABAN F.
PA   (CONL/) CONLEY P B.
PA   (TOPP/) TOPPER J N.
PA   (LAWD/) LAW D.
XX
PI   Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
DR   WPI; 2002-626554/67.
DR   N-PSDB; ABQ98617.
XX
PT   New polypeptide designated ORFX are present in human atherogenic cells
PT   and are useful to prevent and treat ORFX-associated disorders including
PT   cancer, allergy, wound healing or autoimmune, cardiovascular or
PT   inflammatory disease.
XX
PS   Claim 10; SEQ ID NO 848; 78pp; English.
XX
CC   The present invention relates to novel human ORFX polypeptides and their
CC   coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC   were discovered in human atherogenic cells, in particular in platelets
CC   and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC   many other tissues as well. Atherogenic cells are cells which have the
CC   potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC   nucleic acids are useful for treating or preventing a pathological
CC   condition associated with an ORFX-associated disorder, e.g cancer,
CC   cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC   coagulation disorders or inflammatory disorders. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from the USPTO web site at
CC   seqdata.uspto.gov/sequence.html?DocID=2002082206
XX
SQ   Sequence 227 AA;

Query Match   96.6%; Score 1176; DB 5; Length 227;
Best Local Similarity   100.0%; Pred. No. 4.5e-107;
Matches   227; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy   121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy   181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDV 227
              |||||||||||||||||||||||||||||||||||||||||||||||
Db   181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDV 227

RESULT 6
AAY97154
ID   AAY97154 standard; protein; 235 AA.
XX
AC   AAY97154;
XX
DT   04-DEC-2000 (first entry)
XX
DE   Murine regulator of G-protein signaling protein.
XX
KW   RGS; regulators of G-protein signaling; GTPase activating protein; GAP;
KW   G-alpha protein; cell adhesion; chemotaxis; vulnerary; immunosuppressor;
KW   anti-rheumatic; anti-arthritic; anti-diabetic; anti-inflammatory;
KW   cytostatic; hepatotropic; anti-anaemic; modulator; gene therapy.
XX
OS   Mus sp.
XX
FH   Key   Location/Qualifiers
FT   Domain   82..201

```
FT   Misc-difference   83
FT                     /label= RGS_domain
FT   Misc-difference   90
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   100
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   103
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   104
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   107
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   109
FT                     /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference   111
FT                     /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference   112
FT                     /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference   115
FT                     /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference   116
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   138
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   139
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   142
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   143
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   151
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   152
FT                     /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference   154
FT                     /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference   183
FT                     /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference   184
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   187
FT                     /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference   188
FT                     /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference   189
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   191
FT                     /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference   192
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   193
FT                     /note= "forms part of hydrophobic core"
FT   Misc-difference   198
FT                     /note= "forms part of hydrophobic core"
XX
PN   WO20004236-A2.
XX
PD   10-AUG-2000.
XX
PF   04-FEB-2000; 2000WO-US002977.
XX
PR   04-FEB-1999;   99US-00244314.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Hodge MR, Yowe D;
XX
DR   WPI; 2000-532893/48.
DR   N-PSDB; AAA52090.
XX
PT   Novel regulator of G-protein signaling nucleic acids and polypeptides,
PT   useful as diagnostic and investigative tools and to treat G-protein
```

```
PT   signaling disorders.
XX
PS   Claim 8; Page 104; 105pp; English.
XX
CC   The RGS (regulators of G-protein signaling) protein genes, clones
CC   AAH16395 and m1975, were identified in human and murine spleen cDNA
CC   libraries, respectively. Both proteins have unique N- and C-terminal
CC   sequences. The C-terminal location of the RGS domain is consistent with
CC   RGSs known to act as GTPase activating proteins (GAPs) for G-alpha
CC   proteins. G-alpha-i linked receptors support rapid adhesion and directed
CC   migration of leukocytes and other cell types. The novel RGS proteins may
CC   be used to modulate cell adhesion and chemotaxis, e.g. for aiding wound
CC   repair. The RGS proteins, related cDNAs and anti-RGS antibodies are
CC   useful for modulation, diagnosis and treatment of immune and respiratory
CC   disorders
XX
SQ   Sequence 235 AA;

Query Match             84.7%;  Score 1032;  DB 3;  Length 235;
Best Local Similarity   83.8%;  Pred. No. 6.6e-93;
Matches  197;  Conservative   19;  Mismatches   19;  Indels    0;  Gaps    0;

Qy        1 METTLLFFSQINMCESKEKTFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
            |::|:||:||||||||||||||||::||||||||| ||||||||||||||||:|:::|:
Db        1 MDMSLVFFSQLNMCESKEKTFKLMHGSGKEETSIEAKIRAKEKRNRLSLLQRPDFHGE 60

Qy       61 TRSSRGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
            |::||  :|||||||||||||||:||||||||||||| |:||||||||||||||||:|||
Db       61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120

Qy      121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
            ||||  ||||| ||||||||||||| |||||||:|||||||| | :||||||||||||||
Db      121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180

Qy      181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
            |||||:|||:|||||:|||:||:|||||||||||||||||:||:|||:|||||||
Db      181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTNDFQDVKSDVAIWL 235
```

```
XX
DR   WPI; 2002-588886/63.
DR   N-PSDB; AAD42498.
XX
PT   Screening assays for identifying agent that binds to human or mouse RGS
PT   (regulators of G-protein signaling) protein or its variant, or RGS
PT   host cell expressing the RGS protein or its variant, or RGS protein
PT   encoded by homologous DNA sequences, or to host cell expressing
PT   the proteins.
XX
PS   Claim 1; Col 57-60; 42pp; English.
XX
CC   The invention relates to screening assays for identifying an agent that
CC   binds to: human or mouse RGS (regulators of G-protein signalling) protein
CC   or its variant, or RGS protein encoded by homologous DNA sequences; or a
CC   host cell expressing the RGS protein or its variant, or RGS protein
CC   encoded by homologous DNA sequences. The invention is useful for
CC   identifying an agent that binds to human or mouse RGS protein or its
CC   variant, or RGS protein encoded by homologous DNA sequences, or a host
CC   cell expressing the RGS protein or its variant, or RGS protein encoded by
CC   homologous DNA sequences. The agents identified using the invention are
CC   useful for modulating the activity of RGS proteins and thus useful for
CC   treating immune and inflammatory disorders (Grave's disease, allergy,
CC   arthritis), respiratory disorders (asthma, pneumonia, sinusitis),
CC   haematological disorders (haematopoiesis, migration), platelet-associated
CC   disorders (thrombocytopaenia), invasive disorders (leukaemia),
CC   erythrocyte-associated disorders (anaemia), pancreatitis, hepatitis etc.
CC   The present sequence is murine RGS protein, m1975
XX
SQ   Sequence 235 AA;

Query Match            84.7%;  Score 1032;  DB 5;  Length 235;
Best Local Similarity  83.8%;  Pred. No. 6.6e-93;
Matches 197;  Conservative 19;  Mismatches 19;  Indels 0;  Gaps 0;

Qy          1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
              ||:|||||||:|||||||||||||:|:||||||:||||||||||||||||||||
Db          1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLQRPDFHGE 60

Qy         61 TRSSRSGHLAKETRVSPEEAVKWGESPDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
              |:|||::||||||||||||||||||||||:|:|||||||||||||||||||:|:||
Db         61 TQASRSALLAKETRVSPEEAVKWAAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120

Qy        121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSTQPTLHSFDAAQSRV 180
              |||||||||||||||||||:|||||||||||||||||||||||||:||||||||||
Db        121 DFKKCCKBPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180

Qy        181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
              |||||||||:|||:|||:|:||||||||||||||||||:|||||:||||||||
Db        181 YQLMEHDSYKRFLKSETYLHILEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
```

RESULT 8
```
ABP96779
ID   ABP96779 standard; protein; 211 AA.
XX
AC   ABP96779;
XX
DT   05-JUN-2003  (first entry)
XX
DE   Human COPD related protein SEQ ID NO:29.
XX
KW   Human; chronic obstructive pulmonary disease; COPD; chronic lung disease.
XX
OS   Homo sapiens.
XX
PN   WO200297127-A2.
XX
PD   05-DEC-2002.
XX
PF   28-MAY-2002; 2002WO-EP005835.
XX
PR   31-MAY-2001; 2001GB-00013266.
XX
```

```
PA   (FARB ) BAYER AG.
XX
PI   Oellers N, Gehrmann M, Kallabis H, Hall R, Schulze T, Kroegel C;
XX
DR   WPI; 2003-140492/13.
DR   N-PSDB; ACC46750.
XX
PT   Predicting, diagnosing or prognosing chronic lung disease, by detecting a
PT   chronic obstructive pulmonary disease (COPD) gene in a biological sample.
XX
PS   Claim 8; Page 119-120; 214pp; English.
XX
CC   The present invention describes a method for predicting, diagnosing or
CC   prognosing chronic lung disease by detecting a chronic obstructive
CC   pulmonary disease (COPD) gene related polynucleotide (see ACC46750 to
CC   ACC46777, which encode the COPD related proteins in ABP96779 to
CC   ABP96806). The method is useful for predicting, diagnosing or prognosing
CC   chronic lung disease in a biological sample. The COPD genes and proteins
CC   encoded by them from the present invention (I) can be used for treating
CC   or preventing chronic lung disease in a mammal. (I) can be used in an
CC   animal model for determining the efficacy, toxicity, or side effects of
CC   treatment with (I), and determining the mechanism of action of (I).
CC   ACC46778 to ACC46903 represent COPD related PCR primers and probes used
CC   in an example from the present invention
XX
SQ   Sequence 211 AA;

Query Match            34.7%;  Score 422.5;  DB 6;  Length 211;
Best Local Similarity  42.8%;  Pred. No 4.2e-33;
Matches 89;  Conservative 37;  Mismatches 73;  Indels 9;  Gaps 3;

Qy          6 LFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIR---AKEKRNRLSLLVQKPEFHEDTR 62
              |:||: |::|  ||:| :||::|| :|:: |  :|     |: ::|::||:|
Db          5 MFLAVQHDCRPMDKS-----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSTPGKPK 59

Qy         63 SSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDF 122
              |:: |||||:||  :||||||:|||:||||||||||||||||:|||||
Db         60 TGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACEDF 119

Qy        123 KKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSTQPTLHSFDAAQSRVYQ 182
              |::|:  |:: |||||||||||||||:|:||||:||:|||||||:||||||::||:
Db        120 KKTKSPQKLLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRVYS 179

Qy        183 LMEQDSYTRFLKSDIYLDLMEGRPQRPT 210
              ||  :|||:||:|| :: :||||:||:
Db        180 LMENNSYPRFLESEFYQDLCK-KFQITT 206
```

RESULT 9
```
ABP98020
ID   ABP98020 standard; protein; 211 AA.
XX
AC   ABP98020;
XX
DT   11-AUG-2003  (first entry)
XX
DE   Amino acid sequence of regulator polypeptide of G-protein signalling 2.
XX
KW   Human; regulator polypeptide of G-protein signalling; RGS; chromosome 1;
KW   RGS2; asthma; diabetes.
XX
OS   Homo sapiens.
XX
PN   WO200301345-A1.
XX
PD   27-FEB-2003.
XX
PF   16-AUG-2002; 2002WO-EP009173.
PR   16-AUG-2001; 2001US-0312345P.
XX
PA   (FARB ) BAYER AG.
XX
```

```
PI   Koehler RH;
XX
DR   WPI; 2003-278547/27.
XX
PT   New polynucleotide encoding a regulator polypeptide of G-protein
PT   signaling (RGS) and RGS protein, useful in gene therapy to modulate the
PT   activity of RGS in a disease, particularly in asthma or diabetes.
XX
PS   Disclosure; Fig 3; 99pp; English.
XX
CC   The present sequence is a human regulator polypeptide of G-protein
CC   signaling (RGS) 2. The RGS gene is located on chromosome 1. The RGS
CC   polynucleotide and polypeptide are useful for identifying test compounds,
CC   that may act as agonists or antagonists at the receptor site and which
CC   can be regulated to provide therapeutic effects. The polynucleotide,
CC   polypeptide or vectors comprising the polynucleotide, and the reagents
CC   are useful for modulating the activity of RGS in a disease, particularly
CC   in asthma or diabetes
XX
SQ   Sequence 211 AA;

Query Match       34.7%;    Score 422.5; DB 6;  Length 211;
Best Local Similarity  42.8%;  Pred. No. 4.2e-33;
Matches   89;  Conservative  37;  Mismatches  73;  Indels  9;  Gaps  3;

Qy    6  LFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIR---AKEKRNRLSLLVQKPEFHEDTR  62
         :  : .  : ::    |. :|||  ||    : ||| |
Db    5  MFLAVQHDCRPMDKS----AGSGHKSEKREKMKRTLLKDWKTRLSYFLQNSSTPGKPK  59

Qy   63  SSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDF  122
         : : .  ||||.||||| | .: | :|||||: :|| | ||||:|| .||||||||:||||
Db   60  TGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACEDF  119

Qy  123  KKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQ  182
         ||:|.||:: |||||:||| |:.||| |:| |:::: | :|:| |||| : || :|:||:
Db  120  KKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRVYS  179

Qy  183  LMEQDSYTRFLKSDIYLDLMEGRPQRPT  210
         |||::||.||| |:.| |  ||.:| :|
Db  180  LMENNSYPRFLESEFYQDLCK-KPQITT  206


RESULT 10
ADD14019
ID   ADD14019 standard; protein; 211 AA.
XX
AC   ADD14019;
XX
DT   01-JAN-2004 (first entry)
XX
DE   Human src biomarker polypeptide SEQ ID NO:208.
XX
KW   predictor set; protein tyrosine kinase activity modulator;
KW   protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW   gene therapy; drug sensitivity; genetic profile; cancer; human.
XX
OS   Homo sapiens.
XX
PN   WO2003062395-A2.
XX
PD   31-JUL-2003.
XX
PF   17-JAN-2003; 2003WO-US001981.
XX
PR   18-JAN-2002; 2002US-0350061P.
XX
PA   (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI   Huang F, Fairchild CR, Lee FY, Shaw P;
XX
DR   WPI; 2003-636735/60.
DR   N-PSDB; ADD14613.
XX
PT   New polynucleotides and polypeptides for predicting the activity of
PT   compounds that interact with protein tyrosine kinases and/or protein
PT   tyrosine kinase pathways.
XX
PS   Claim 10; SEQ ID NO 208; 139pp; English.
XX
CC   The present invention describes a predictor set comprising a plurality of
CC   polynucleotides or polypeptides whose expression pattern is predictive of
CC   the response of cells to treatment with a compound that modulates protein
CC   tyrosine kinase activity or members of the protein tyrosine kinase
CC   pathway. Also described: (1) predicting whether a compound is capable of
CC   modulating the activity of cells, comprising obtaining a sample of cells,
CC   determining whether the cells express a plurality of markers, and
CC   correlating the expression of the markers to the compound's ability to
CC   modulate the activity of the cells; (2) a plurality of cell lines for
CC   identifying polynucleotides and polypeptides whose expression levels
CC   correlate with compound sensitivity or resistance of cells associated
CC   with a disease state; and (3) identifying polynucleotides and
CC   polypeptides that predict compound sensitivity or resistance of cells
CC   associated with a disease state, comprising subjecting the plurality of
CC   cell lines to one or more compounds, analysing the expression pattern of
CC   a microarray of polynucleotides or polypeptides, and selecting
CC   polynucleotides or polypeptides that predict the sensitivity or
CC   resistance of cells associated with a disease state by using the
CC   expression pattern of the microarray. The polynucleotides and
CC   polypeptides have cytostatic activities, and can be used in gene therapy.
CC   The polynucleotides and polypeptides are useful in predicting the
CC   activity of compounds that interact with protein tyrosine kinases and/or
CC   protein tyrosine kinase pathways. These may be used in determining drug
CC   sensitivity in patients to allow the development of individualized
CC   genetic profiles which aid in treating diseases and disorders (e.g.
CC   cancer) based on patient response at a molecular level. The present
CC   sequence is used in the exemplification of the present invention.
XX
SQ   Sequence 211 AA;

Query Match       34.7%;    Score 422.5; DB 7;  Length 211;
Best Local Similarity  42.8%;  Pred. No. 4.2e-33;
Matches   89;  Conservative  37;  Mismatches  73;  Indels  9;  Gaps  3;

Qy    6  LFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTR  62
         :  : .  : ::    |. :|||  ||    : ||| |
Db    5  MFLAVQHDCRPMDKS----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSSTPGKPK  59

Qy   63  SSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDF  122
         : : .  ||||.||||| | .: | :|||||: :|| | ||||:|| .||||||||:||||
Db   60  TGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACEDF  119

Qy  123  KKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQ  182
         ||:|.||:: |||||:||| |:.||| |:| |:::: | :|:| |||| : || :|:||:
Db  120  KKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRVYS  179

Qy  183  LMEQDSYTRFLKSDIYLDLMEGRPQRPT  210
         |||::||.||| |:.| |  ||.:| :|
Db  180  LMENNSYPRFLESEFYQDLCK-KPQITT  206


RESULT 11
AAB56929
ID   AAB56929 standard; protein; 220 AA.
XX
AC   AAB56929;
XX
DT   13-MAR-2001 (first entry)
XX
DE   Human prostate cancer antigen protein sequence SEQ ID NO:1507.
XX
KW   Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW   neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW   vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW   antibacterial; gene therapy; neural; immune; reproductive; renal;
KW   gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW   wound; infectious disease.
```

XX
OS   Homo sapiens.
XX
PN   WO200055174-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US005988.
XX
PR   12-MAR-1999; 99US-0124270P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (ROSE/) ROSEN C A.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2000-587513/55.
DR   N-PSDB; AAF16132.
XX
PT   Prostate cancer associated gene sequences, referred to as prostate cancer
PT   antigens, useful for treatment, prevention, and diagnosis of disorders
PT   such as prostate cancer.
XX
PS   Claim 11; Page 1946-1947; 2338pp; English.
XX
CC   AAF15566 to AAF16505 encode the human prostate cancer associated
CC   proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC   The prostate cancer antigens can have neuroprotective, cytostatic,
CC   cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC   nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC   and can be used in gene therapy. The prostate cancer antigen
CC   polynucleotides may be used for detection of prostate cancer, chromosome
CC   identification, as chromosome markers, and for numerous other diagnostic
CC   or research purposes. The prostate cancer antigens may be used to treat
CC   disorders such as neural, immune, muscular, reproductive,
CC   gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC   disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC   AAB57303 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 220 AA;

Query Match   34.7%; Score 422.5; DB 3; Length 220;
Best Local Similarity   42.8%; Pred. No. 4.5e-33;
Matches   89; Conservative   37; Mismatches   73; Indels   9; Gaps   3;

Qy   6 LFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTR 62
         |:: :||||| : :|| | | || :|||||:||||||
Db   14 MFLAVQHDCRPMDKS----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSSTPGKPK 68

Qy   63 SSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDF 122
         : : :||||| :|:||:|| ||:| : | |||:|:|| |||||
Db   69 TGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENEFWLACEDF 128

Qy 123 KKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQ 182
         :: : |||: ||::| | | | ||:|: ||:|||||
Db 129 KKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRVYS 188

Qy 183 LMEQDSYTRFLKSDIYLDLMEGRPQRPT 210
         |||:|||:||: |: ||:
Db 189 LMENNSYTPRFLESEFYQDLCK-KPQITT 215

XX
DE   Human regulator of G-protein signalling.

RESULT 12
AAW62075
ID   AAW62075 standard; protein; 181 AA.
XX
AC   AAW62075;
XX
DT   14-SEP-1998 (first entry)
XX

KW   Human; regulator; G-protein signalling; HRGS; cancer; inflammation;
KW   hypertension; cardiovascular shock; arrhythmias; asthma.
XX
OS   Homo sapiens.
XX
PN   WO9820128-A1.
XX
PD   14-MAY-1998.
XX
PF   06-NOV-1997;   97WO-US018476.
XX
PR   08-NOV-1996;   96US-00748483.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Hillman JL, Goli SK;
XX
DR   WPI; 1998-286944/25.
DR   N-PSDB; AAV38084.
XX
PT   Regulator of G-protein signalling - used to develop products for treating
PT   e.g. cancer, inflammation, hypertension, cardiovascular shock,
PT   arrhythmias or asthma.
XX
PS   Claim 1; Fig 1; 66pp; English.
XX
CC   The present sequence represents human regulator of G-protein signalling
CC   (HRGS). The HRGS regulates G-protein signalling in cancer cells and may
CC   be useful in the treatment of any cancer, especially cancers of the brain
CC   and thyroid. Products of the present invention can also be used for
CC   treating other conditions associated with uncontrolled cell signalling
CC   such as inflammation. The products can also be used to modulate HRGS
CC   activity in response to disorders involving the sympathetic nervous
CC   system including hypertension, cardiovascular shock, arrhythmias and
CC   asthma. The products can also be used for detection, diagnosis and drug
CC   screening
XX
SQ   Sequence 181 AA;

Query Match   34.6%; Score 421.5; DB 2; Length 181;
Best Local Similarity   49.1%; Pred. No. 4.3e-33;
Matches   82; Conservative   37; Mismatches   45; Indels   3; Gaps   1;

Qy   40 RAKEKRNRLSLLVQKPEFHED--TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG 96
         |||:|| : :|||:| |:|| :|:|| |||||||||||||||||||||||||||||
Db   15 RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG 74

Qy   97 LEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
         :|:|||||||||:|||||||||||| : :|||||:|||||||||||||
Db   75 LASFKSFLKSEFSEENLEFWIACEDYKKIKSPAKMAEKAKQIYEEFIQEAPKEVNIDHF 134

Qy 157 TKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLME 203
         |||:|||: :|:||||||||||||||| |||||:||||||||||
Db 135 TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK 181

XX

RESULT 13
AAY53931
ID   AAY53931 standard; protein; 181 AA.
XX
AC   AAY53931;
XX
DT   13-MAR-2000 (first entry)
XX
DE   A human regulator of G protein signalling RGS5 protein.
XX
KW   Activator of G protein signalling; AGS; ras-related G protein;
KW   GTP hydrolysis; G protein activity; pheromone response pathway;
KW   G protein-coupled signal transduction; G-gamma selectivity;
KW   cellular signal transduction; regulator of G protein signalling; RGS5.
XX
OS   Homo sapiens.
XX

```
PN   WO9958670-A1.
XX
PD   18-NOV-1999.
XX
PF   07-MAY-1999;   99WO-US010151.
XX
PR   08-MAY-1998;   98US-0084842P.
PR   07-OCT-1998;   98US-0103355P.
XX
PA   (CADU-) CADUS PHARM CORP.
XX
PI   Cismowski M, Duzic E;
XX
DR   WPI; 2000-072337/06.
DR   N-PSDB; AA236910.
XX
PT   A new activator of G protein signalling used to treat disorders
     characterized by an aberrant AGS protein activity.
XX
PS   Claim 74; Page 141-141; 162pp; English.
XX
CC   The present sequence represents a regulator of G protein signalling (AGS)
CC   protein, RGS5. The specification also describes an activator of G protein
CC   signalling (AGS) protein. The AGS cDNA sequence was isolated from a human
CC   liver cDNA library. The AGS protein exhibits homology to ras-related G
CC   proteins, and contains alterations in conserved amino acids consistent
CC   with a deficiency in GTP hydrolysis activity. AGS stimulates G protein
CC   activity, G protein-coupled signal transduction and the pheromone
CC   response pathway in a receptor-independent manner. The AGS protein also
CC   shows G-gamma selectivity, as measured by growth assays in yeast
CC   expressing various mammalian G-gamma constructs, and tissue-specific
CC   expression, as measured by Northern blot analysis. The AGS protein can be
CC   used to screen for compounds that modulate cellular signal transduction.
CC   The protein is used to treat disorders characterized by an aberrant AGS
CC   protein activity or AGS nucleic acid expression
XX
SQ   Sequence 181 AA;

Query Match             34.6%;  Score 421.5;  DB 3;  Length 181;
Best Local Similarity   49.1%;  Pred. No. 4.3e-33;
Matches  82;  Conservative  37;  Mismatches  45;  Indels  3;  Gaps  1;

Qy         40 RAKEKRNRLSLLVQKPEFHED--TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG 96
              ||||:||:|:||  :|::|: |  |||:|:|  | ||:|:|:|   |||||
Db         15 RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG 74

Qy         97 LEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
              |:|:|||||||||| :|:  |||:|:| | ||:|||||:||||:||||||:||||:|||:
Db         75 LASFKSFLKSEFSEENLEFWIACEDYKKIKSPAKMAEKAKQIYEEFIQTERPKEVNIDHF 134

Qy        157 TKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLME 203
              ||:||:|:| ||| |:||||:| :||::::|:|||||:|:|:
Db        135 TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK 181


RESULT 14
ABP62144
ID   ABP62144 standard; protein; 190 AA.
XX
AC   ABP62144;
XX
DT   12-NOV-2002 (first entry)
XX
DE   Human secreted protein SEQ ID NO 197.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
```

```
OS   Homo sapiens.
XX
PN   WO200257420-A2.
XX
PD   25-JUL-2002.
XX
PF   17-JAN-2002;   2002WO-US001109.
XX
PR   18-JAN-2001;   2001US-0262066P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Moore PA, Ruben SM, Lafleur DW, Shi Y, Rosen CA, Olsen H;
PI   Ebner R, Brewer LA;
XX
DR   WPI; 2002-599716/64.
XX
PT   New polynucleotides and polypeptides useful for diagnosing, prognosing,
PT   treating or preventing e.g. neurodegenerative, central nervous system,
PT   autoimmune, respiratory, reproductive, or inflammatory diseases or
PT   disorders.
XX
PS   Claim 11; Page 71; 785pp; English.
XX
CC   The invention relates to novel genes (ABQ92553-ABQ92607) and proteins
CC   (ABP62013-ABP62153) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC   breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
CC   infections
XX
SQ   Sequence 190 AA;

Query Match             34.6%;  Score 421.5;  DB 5;  Length 190;
Best Local Similarity   49.1%;  Pred. No. 4.6e-33;
Matches  82;  Conservative  37;  Mismatches  45;  Indels  3;  Gaps  1;

Qy         40 RAKEKRNRLSLLVQKPEFHED---TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG 96
              ||||:||:|:||  :|::|:  |  |||:|:|  | ||:|:|:|   |||||
Db         24 RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG 83

Qy         97 LEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
              |:|:|||||||||| :|:  |||:|:| | ||:|||||:||||:||||||:||||:|||:
Db         84 LASFKSFLKSEFSEENLEFWIACEDYKKIKSPAKMAEKAKQIYEEFIQTERPKEVNIDHF 143

Qy        157 TKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLME 203
              ||:||:|:| ||| |:||||:| :||::::|:|||||:|:|:
Db        144 TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK 190


RESULT 15
AAU10747
ID   AAU10747 standard; protein; 81 AA.
XX
AC   AAU10747;
XX
DT   12-MAR-2002 (first entry)
XX
DE   Amino acid sequence of partial human platelet RGS domain.
XX
KW   Human; regulator of G protein signalling; RGS18; arterial thrombosis;
KW   platelet activation dysfunction; myocardial infarction; stroke;
KW   coronary artery disease; cerebrovascular disease; unstable angina;
KW   deep vein thrombosis; systemic thromboembolism; anti-coagulant;
XX
```

```
KW   invasive cardiac procedure.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          2..81
FT                   /note= "Encoded by sequence in AAS18334"
XX
PN   WO200183514-A2.
XX
PD   08-NOV-2001.
XX
PF   26-APR-2001; 2001WO-US013540.
XX
PR   28-APR-2000; 2000US-0200786P.
PR   02-AUG-2000; 2000GB-0018833.
XX
PA   (AVET ) AVENTIS PHARM PROD INC.
XX
PI   Murray DL, Gagnon AW;
XX
DR   WPI; 2002-055453/07.
DR   N-PSDB; AAS18334.
XX
PT   Isolated regulator of G protein signaling polypeptide, useful for
PT   prevention/treatment of platelet activation dysfunction, such as arterial
PT   thrombosis, myocardial infarction, coronary artery disease and stroke.
XX
PS   Claim 31; Page 115; 127pp; English.
XX
CC   The present invention relates to the isolation of polynucleotide
CC   sequences, that encode a novel regulator of G protein signalling (RGS)
CC   polypeptide, RGS18, from human platelets. The invention also provides
CC   nucleotide primers and probes specific for an RGS18 nucleic acid. The
CC   sequences of the invention are useful for the manufacture of a medicament
CC   for the prevention or treatment of a platelet activation dysfunction,
CC   such as arterial thrombosis, myocardial infarction, coronary artery
CC   disease, stroke, cerebrovascular disease, unstable angina, deep vein
CC   thrombosis, systemic thromboembolism, and in invasive cardiac procedures
CC   for anti-coagulant purposes. The sequences of the invention can also be
CC   used for the detection of RGS18 nucleic acids and RGS18 polypeptides. The
CC   present sequence represents a partial human platelet RGS domain
XX
SQ   Sequence 81 AA;

Query Match             33.9%;  Score 413;  DB 5;  Length 81;
Best Local Similarity   97.5%;  Pred. No. 9.8e-33;
Matches   78;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    110 EENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPT 169
              ||:||||:|||||||||||||||:|||||||||||||||||||||||:|||||||||||
Db      2 EENLEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPT 61

Qy    170 LHSFDAAQSRVYQLMEQDSY 189
              |||||||||||||||  |||
Db     62 LHSFDAAQSRVYQLMENDSY 81
```

Search completed: August 20, 2004, 16:44:14
Job time : 81 secs

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:         August 20, 2004, 16:40:58 ; Search time 18.5 Seconds
                                             (without alignments)
                                             655.789 Million cell updates/sec

Title:          US-09-894-749-2
Perfect score:  1218
Sequence:       1 METTLLFFSQINMCESKEKT.........SRSFTCNEFQDVQSDVAIWL 235

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters:    389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Issued_Patents AA:*
            1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
            2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
            3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
            4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
            5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
            6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1218	100.0	235	3	US-09-244-314-2	Sequence 2, Appli
2	1218	100.0	235	4	US-09-498-959-2	Sequence 2, Appli
3	1032	84.7	235	3	US-09-244-314-4	Sequence 4, Appli
4	1032	84.7	235	4	US-09-498-959-4	Sequence 4, Appli
5	422.5	34.7	211	2	US-08-748-483-4	Sequence 4, Appli
6	421.5	34.6	181	2	US-08-748-483-1	Sequence 1, Appli
7	408.5	33.5	205	2	US-08-829-110-6	Sequence 6, Appli
8	408.5	33.5	205	2	US-08-748-483-5	Sequence 5, Appli
9	380	31.2	202	1	US-08-274-318-2	Sequence 2, Appli
10	380	31.2	202	2	US-08-463-081B-2	Sequence 2, Appli
11	380	31.2	202	2	US-08-461-379A-2	Sequence 2, Appli
12	380	31.2	202	2	US-08-462-390B-2	Sequence 2, Appli
13	380	31.2	202	2	US-08-754-108-2	Sequence 2, Appli
14	380	31.2	202	3	US-08-870-815-4	Sequence 4, Appli
15	380	31.2	202	3	US-08-463-074S-2	Sequence 2, Appli
16	380	31.2	202	3	US-08-465-585C-2	Sequence 2, Appli
17	380	31.2	202	3	US-08-652-446-2	Sequence 2, Appli
18	380	31.2	202	3	US-08-949-004-4	Sequence 4, Appli
19	369	30.3	120	4	US-08-890-865A-11	Sequence 11, Appl
20	361	29.6	201	2	US-08-726-228-2	Sequence 2, Appli
21	361	29.6	201	3	US-08-870-815-2	Sequence 2, Appli
22	361	29.6	201	3	US-08-949-004-2	Sequence 2, Appli
23	359	29.5	120	4	US-08-890-865A-13	Sequence 13, Appli
24	359.5	29.4	121	2	US-08-588-258B-32	Sequence 32, Appl
25	358.5	29.4	121	3	US-08-460-505-32	Sequence 32, Appl
26	358.5	29.4	121	5	PCT-US96-08295-32	Sequence 32, Appl
27	358	29.4	120	4	US-08-890-865A-14	Sequence 14, Appl
28	357	29.3	120	4	US-08-890-865A-12	Sequence 12, Appl
29	353	29.0	196	2	US-08-829-110-5	Sequence 5, Appli
30	353	29.0	196	2	US-08-748-483-3	Sequence 3, Appl
31	353	29.0	196	4	US-09-702-705-339	Sequence 339, App
32	353	29.0	196	4	US-09-736-457-339	Sequence 339, App
33	353	29.0	196	4	US-09-614-124B-339	Sequence 339, App
34	353	29.0	196	4	US-09-671-325-339	Sequence 339, App
35	335.5	27.5	119	2	US-09-589-184-339	Sequence 339, App
36	335.5	27.5	119	2	US-08-588-258B-31	Sequence 31, Appl
37	335.5	27.0	119	5	PCT-US96-08295-31	Sequence 31, Appl
38	328.5	27.0	119	4	US-08-890-865A-15	Sequence 15, Appl
39	306	25.1	217	4	US-09-976-594-51	Sequence 51, Appl
40	290.5	23.9	119	3	US-08-588-258B-30	Sequence 30, Appl
41	290.5	23.9	119	3	US-08-460-505-30	Sequence 30, Appl
42	290.5	23.9	119	5	PCT-US96-08295-30	Sequence 30, Appl
43	287	23.6	159	2	US-08-829-110-1	Sequence 1, Appl
44	283	23.2	120	4	US-08-890-865A-16	Sequence 16, Appl

ALIGNMENTS

```
RESULT 1
US-09-244-314-2
; Sequence 2, Application US/09244314
; Patent No. 6274362
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; APPLICANT: Yowe, David
; TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
; FILE REFERENCE: 5800-19, 03800/17680
; CURRENT APPLICATION NUMBER: US/09/244,314
; CURRENT FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 235
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-244-314-2

Query Match             100.0%;  Score 1218;  DB 3;  Length 235;
Best Local Similarity   100.0%;  Pred. No. 3.6e-116;
Matches  235;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1 METTLLFFSQINMCESKEKTFKLIHGSGKEETSKEAKIRAKEKRNLSLLVQKPEFHED 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 METTLLFFSQINMCESKEKTFKLIHGSGKEETSKEAKIRAKEKRNLSLLVQKPEFHED 60

Qy   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVIINSIIQPTLHSFDAAQSRV 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVIINSIIQPTLHSFDAAQSRV 180

Qy  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRSRSFTCNEFQDVQSDVAIWL 235
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRSRSFTCNEFQDVQSDVAIWL 235


RESULT 2
US-09-498-959-2
; Sequence 2, Application US/09498959
; Patent No. 6410240
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; APPLICANT: Yowe, David
; TITLE OF INVENTION: RGS-Containing Molecules and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 5800-19A
```

```
;    CURRENT APPLICATION NUMBER: US/09/498,959
;    CURRENT FILING DATE: 2000-02-04
;    EARLIER APPLICATION NUMBER: 09/244,314
;    EARLIER FILING DATE: 1999-02-04
;    NUMBER OF SEQ ID NOS: 12
;    SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;    LENGTH: 235
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-498-959-2

  Query Match             100.0%;  Score 1218;  DB 4;  Length 235;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-116;
  Matches  235;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy         61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy        121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy        181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235

RESULT 3
US-09-244-314-4
; Sequence 4, Application US/09244314
; Patent No. 6274362
; GENERAL INFORMATION:
;    APPLICANT: Hodge, Martin R.
;    APPLICANT: Yowe, David
;    TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
;    FILE REFERENCE: 5800-19, 035800/174680
;    CURRENT APPLICATION NUMBER: US/09/244,314
;    CURRENT FILING DATE: 1999-02-04
;    NUMBER OF SEQ ID NOS: 4
;    SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;    LENGTH: 235
;    TYPE: PRT
;    ORGANISM: Mus sp.
US-09-244-314-4

  Query Match              84.7%;  Score 1032;  DB 3;  Length 235;
  Best Local Similarity    83.8%;  Pred. No. 3.2e-97;
  Matches  197;  Conservative   19;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
              | :  |:|||| ||||||||||||||:| |||||| ||||||||||||||||| :|| ::
Db          1 MDMSLVFFSQLNMCESKEKTFFKLMEGSGKEETSIEAKIRAKEKRNRLSLLQRPDFHGE 60

Qy         61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
              | :|||: ||||||||||||||| :|||||||||||: ||||||||||||||||||:||
Db         61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120

Qy        121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
              ||||  :||:|||||||||||||| ||||||:|||||||| :: ||:||||||||:|||||
Db        121 DFKKCKEPQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180

Qy        181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
              |||||:||| :|||| | :| :||||||||||||||||||| :||||| ||||||
Db        181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
```

```
RESULT 4
US-09-498-959-4
; Sequence 4, Application US/09498959
; Patent No. 6410240
; GENERAL INFORMATION:
;    APPLICANT: Hodge, Martin R.
;    APPLICANT: Yowe, David
;    TITLE OF INVENTION: RGS-Containing Molecules and Uses
;    TITLE OF INVENTION: Thereof
;    FILE REFERENCE: 5800-19A
;    CURRENT APPLICATION NUMBER: US/09/498,959
;    CURRENT FILING DATE: 2000-02-04
;    EARLIER APPLICATION NUMBER: 09/244,314
;    EARLIER FILING DATE: 1999-02-04
;    NUMBER OF SEQ ID NOS: 12
;    SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;    LENGTH: 235
;    TYPE: PRT
;    ORGANISM: Mus sp.
US-09-498-959-4

  Query Match              84.7%;  Score 1032;  DB 4;  Length 235;
  Best Local Similarity    83.8%;  Pred. No. 3.2e-97;
  Matches  197;  Conservative   19;  Mismatches   19;  Indels    0;  Gaps    0;

Qy          1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
              | :  |:|||| ||||||||||||||:| |||||| ||||||||||||||||| :|| ::
Db          1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLQRPDFHGE 60

Qy         61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
              | :|||: ||||||||||||||| :|||||||||||: ||||||||||||||||||:||
Db         61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120

Qy        121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
              ||||  :||:|||||||||||||| ||||||:|||||||| :: ||:||||||||:|||||
Db        121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180

Qy        181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
              |||||:||| :|||| | :| :||||||||||||||||||| :||||| ||||||
Db        181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235

RESULT 5
US-08-748-483-4
; Sequence 4, Application US/08748483
; Patent No. 5995314
; GENERAL INFORMATION:
;    APPLICANT: Hillman, Jennifer L.
;    APPLICANT: Goli, Surya K.
;    TITLE OF INVENTION: A NOVEL REGULATOR OF CELL SIGNALLING
;    NUMBER OF SEQUENCES: 5
;    CORRESPONDENCE ADDRESS:
;    ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;    STREET: 3174 Porter Drive
;    CITY: Palo Alto
;    STATE: CA
;    COUNTRY: US
;    ZIP: 94304
;    COMPUTER READABLE FORM:
;    MEDIUM TYPE: Diskette
;    COMPUTER: IBM Compatible
;    OPERATING SYSTEM: DOS
;    SOFTWARE: FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;    APPLICATION NUMBER: US/08/748,483
;    FILING DATE: Herewith
;    CLASSIFICATION: 530
;    PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:
;    FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;    NAME: Billings, Lucy J.
```

```
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0157 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 211 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 292037
US-08-748-483-4

Query Match          34.7%;   Score 422.5; DB 2; Length 211;
Best Local Similarity  42.8%;   Pred. No. 3.4e-35;
Matches   89;  Conservative 37;  Mismatches 73;  Indels 9;  Gaps 3;

Qy   6 LFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIR---AKEKRNRLSLLVQKPEFHEDTR  62
         :|: .|||:  :|:  :||  |:|| :|||    |||:| |
Db   5 MFLAVQHDCRPMDKS-----AGSGHKSEKRFKMKRTLLKDWKTRLSYFLQNSSTPGKPK  59

Qy  63 SSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSSEENIEFWIACEDF 122
        : :|||||  |-|||:||| |||| |:|||||:||||||
Db  60 TGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACEDF 119

Qy 123 KKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQ 182
        ||-||||:  |||:|||||||||:|||:||:|  : .||||
Db 120 KKTKSPQKJSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRVYS 179

Qy 183 LMEQDSYTRFLKSDIYLDLMEGRPQRPT 210
        ||.||||||||||||:.||:|
Db 180 LMENNSYPRFLESEFYQDLICK-KPQITT 206

RESULT 6
US-08-748-483-1
; Sequence 1, Application US/08748483
; Patent No. 5955314
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: A NOVEL REGULATOR OF CELL SIGNALLING
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/748,483
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0157 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
```

```
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 181 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: 57362
US-08-748-483-1

Query Match          34.6%;   Score 421.5; DB 2; Length 181;
Best Local Similarity  49.1%;   Pred. No. 3.5e-35;
Matches   82;  Conservative 37;  Mismatches 45;  Indels 3;  Gaps 1;

Qy  40 RAKEKRNRLSLLVQKPFHED--TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG  96
        |||:||  :.:|||:||  :.|  ||:: |: :||| :.|:| |||| |:||||||
Db  15 RAKEIKIKLGIILQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG  74

Qy  97 LEAFTRFLKTEFSSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
        |||:||||  :||||:||||||| |:| ||: ::|:.||||||:||:|
Db  75 LASFKSFIKSEFSEENLEFWIACEDYKKIKSPAKMABKAKQIYEEFIQTEAPKEVNIDHF 134

Qy 157 TKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLME 203
        ||:|  :|||:||||||||||| :.||||||||||| |:
Db 135 TKDITMNLVEPSLSSFDMAQKRIHALMEKDSLPRVRSEFYQELIK 181

RESULT 7
US-08-829-110-6
; Sequence 6, Application US/08829110
; Patent No. 5882890
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL REGULATORS OF G-PROTEIN
TITLE OF INVENTION: SIGNALING
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/829,110
FILING DATE: Filed Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0259 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 205 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
```

;
US-08-829-110-6

Query Match   33.5%; Score 408.5; DB 2; Length 205;
Best Local Similarity 49.1%; Pred. No. 8.8e-34;
Matches   82; Conservative 28; Mismatches 48; Indels   9; Gaps.   2;

Qy   41 AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKE-----TRVSPEEAVKWGESFDKLLSHRD   95
Db   16 AKDMKHRLGFLLQK----SDSCEHNSSHNKKDKVVICQRVSQEEVKKWAESLENLISHEC   71

Qy   96 GLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDF 155
Db   72 GLAAFKAFLKSEYSEENIDFWISCEEYKKIKSPSKLLSPKAKKIYNEFISVQATKEVNLDS 131

Qy   156 HTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLM 202
Db   132 CTREETSRNMLEPTITCFDEAQKKIFNLMEKDSYRRFLKSRFYLDLV 178

RESULT 8
US-08-748-483-5
; Sequence 5, Application US/08748483
; GENERAL INFORMATION:
;   APPLICANT: Hillman, Jennifer L.
;   APPLICANT: Goli, Surya K.
;   TITLE OF INVENTION: A NOVEL REGULATOR OF CELL SIGNALLING
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 Porter Drive
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: US
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/748,483
;   FILING DATE: Herewith
;   CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Billings, Lucy J.
;   REGISTRATION NUMBER: 36,749
;   REFERENCE/DOCKET NUMBER: PF-0157 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-855-0555
;   TELEFAX: 415-845-4166
;   TELEX:
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 205 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   LIBRARY: GenBank
;   CLONE: 1216373
US-08-748-483-5

Query Match   33.5%; Score 408.5; DB 2; Length 205;
Best Local Similarity 49.1%; Pred. No. 8.8e-34;
Matches   82; Conservative 28; Mismatches 48; Indels   9; Gaps   2;

Qy   41 AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKE-----TRVSPEEAVKWGESFDKLLSHRD   95

Db   16 AKDMKHRLGFLLQK----SDSCEHNSSHNKKDKVVICQRVSQEEVKKWAESLENLISHEC   71

Qy   96 GLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDF 155
Db   72 GLAAFKAFLKSEYSEENIDFWISCEEYKKIKSPSKLLSPKAKKIYNEFISVQATKEVNLDS 131

Qy   156 HTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLM 202
Db   132 CTREETSRNMLEPTITCFDEAQKKIFNLMEKDSYRRFLKSRFYLDLV 178

RESULT 9
US-08-274-318-2
; Sequence 2, Application US/08274318
; Patent No. 5667987
; GENERAL INFORMATION:
;   APPLICANT: Buckbinder, Leonard
;   APPLICANT: Talbott, Randy
;   APPLICANT: Seizinger, Bernd R.
;   APPLICANT: Kley, Nikolai
;   TITLE OF INVENTION: No. 5667987e1 p53 Response Genes
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Bristol-Myers Squibb Company
;   STREET: 3005 First Avenue
;   CITY: Seattle
;   STATE: Washington
;   COUNTRY: USA
;   ZIP: 98121
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/274,318
;   FILING DATE: 11-JUL-1994
;   CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
;   NAME: Bogden, James M.
;   REGISTRATION NUMBER: 32,962
;   REFERENCE/DOCKET NUMBER: ON0127-
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 206-727-3688
;   TELEFAX: 206-727-3601
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 202 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-274-318-2

Query Match   31.2%; Score 380; DB 1; Length 202;
Best Local Similarity 48.5%; Pred. No. 6.9e-31;
Matches   81; Conservative 28; Mismatches 52; Indels   6; Gaps   2;

Qy   40 RAKEKRNRLSLLVQKPEFHEDTRSS-----RSGHLAKETRVSPEEAVKWGESFDKLLSHR   94
Db   15 RAKEFKTRLGIFLHKSELGCDTGSTGKSEWGSKH-SKENRNFSEDVLGWRESFDLLLSSK   73

Qy   95 DGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLD 154
Db   74 NGVAAFHAFLKTEFSSENLEFWLACEFKKIRSATKLASRAHQIFEEFICSEAPKEVNID 133

Qy   155 FHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDL 201
Db   134 HETRELTRMNLQTATATCFDAAQGKTRTLMEKDSYPRFLKSPAYRDL 180

RESULT 10
US-08-463-081B-2
; Sequence 2, Application US/08463081B

Patent No. 5871960 5837487
GENERAL INFORMATION:
APPLICANT: Smith, Kendall A. & Beadling, Carol
TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
TITLE OF INVENTION: Vector and Transformed Cell Thereof, and Expression Thereof
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
STREET: 444 South Flower St. - Suite 1900
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,081B
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/104,736
FILING DATE: 10-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/796,066
FILING DATE: 20-NOV-91
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Amzel, Ph.D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38150 (DART-060)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 622-7700
TELEFAX: (213) 489-4210
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 202 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-081B-2

Query Match   31.2%;   Score 380; DB 2;   Length 202;
Best Local Similarity 48.5%;   Pred. No. 6.9e-31;
Matches 81; Conservative 28; Mismatches 52; Indels 6; Gaps 2;

Qy   40 RAKEKRNRLSLLVQKPEFHEDTRSS-----RSGHLAKETRVSPEEAVKWGESFDKLLSHR 94
         |||| || |:|||: |: | || | ||:| :| |||||: |||
Db   15 RAKEFKTRLGIFLHKSELGCDTGSTGKFEWGSKH-SKENRNFSEDVLGWRESFDLLLSSK 73

Qy   95 DGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLD 154
         :|: || ||||||||||:|||:|| |:||| : |:| ::| :|| :|| :||||||:|
Db   74 NGVAAFHAFLKTEFSEENLEFWLACEEFKKIRSATKLASRAHQIFEEFICSEAPKEVNID 133

Qy 155 FHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDL 201
         | |:|:   :|: | : ||||:| | ||||:| |||| | ||
Db 134 HETRELTRMNLQTATATCFDAAQGKTRTLMEKDSYPRFLKSPAYRDL 180

RESULT 11
US-08-461-379A-2
; Sequence 2, Application US/08461379A
; Patent No. 5871961
; GENERAL INFORMATION:
  APPLICANT: Smith, Kendall A. & Beadling, Carol
  TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
  TITLE OF INVENTION: Vector and Transformed Cell Thereof, and
  TITLE OF INVENTION: Expression Thereof
  NUMBER OF SEQUENCES: 35
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Ratner & Prestia

ADDRESSE: (B) STREET:One Westlakes-Berwyn
CITY: Valley Forge
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,379A
FILING DATE: 5-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/330,108; 08/104,736
APPLICATION NUMBER: & 07/796,066
FILING DATE: 27-OCT-1994; 10-AUG-1993 & 20-NOV-91
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Amzel, Ph.D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: DART-070
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610)470-0700
TELEFAX: (610)470-0701
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 202 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-379A-2

Query Match   31.2%;   Score 380; DB 2;   Length 202;
Best Local Similarity 48.5%;   Pred. No. 6.9e-31;
Matches 81; Conservative 28; Mismatches 52; Indels 6; Gaps 2;

Qy   40 RAKEKRNRLSLLVQKPEFHEDTRSS-----RSGHLAKETRVSPEEAVKWGESFDKLLSHR 94
         |||| || |:|||: |: | || | ||:| :| |||||: |||
Db   15 RAKEFKTRLGIFLHKSELGCDTGSTGKFEWGSKH-SKENRNFSEDVLGWRESFDLLLSSK 73

Qy   95 DGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLD 154
         :|: || ||||||||||:|||:|| |:||| : |:| ::| :|| :|| :||||||:|
Db   74 NGVAAFHAFLKTEFSEENLEFWLACEEFKKIRSATKLASRAHQIFEEFICSEAPKEVNID 133

Qy 155 FHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDL 201
         | |:|:   :|: | : ||||:| | ||||:| |||| | ||
Db 134 HETRELTRMNLQTATATCFDAAQGKTRTLMEKDSYPRFLKSPAYRDL 180

RESULT 12
US-08-462-390B-2
; Sequence 2, Application US/08462390B
; Patent No. 5882894
; GENERAL INFORMATION:
  APPLICANT: Smith, K. A.; & Beadling, C.
  TITLE OF INVENTION: Nucleic Acids Encoding CR8 Polypeptide, Vector and
  TITLE OF INVENTION: Transformed Cell Thereof, and Expression Thereof
  NUMBER OF SEQUENCES: 35
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Ratner & Prestia   (B) STREET:One Westlakes-Berwyn
  ADDRESSEE: Valley Forge
  CITY: Valley Forge
  STATE: Pennsylvania
  COUNTRY: USA
  ZIP: 19482
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/462,390B

;   FILING DATE: 5-JUNE-1995
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: USSN 08/330,108
;   FILING DATE: 27-OCT-1994
;   APPLICATION NUMBER: USSN 08/104,736
;   FILING DATE: 10-AUG-1993
;   APPLICATION NUMBER: USSN 07/796,066
;   FILING DATE: 20-NOV-91
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Viviana Amzel, Ph. D.
;   REGISTRATION NUMBER: 30,930
;   REFERENCE/DOCKET NUMBER: DART-040
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (610)407-0700
;   TELEFAX: (610)407-0701
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 202 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-462-390B-2

Query Match   31.2%;   Score 380;   DB 2;   Length 202;
Best Local Similarity   48.5%;   Pred. No. 6.9e-31;
Matches   81;   Conservative   28;   Mismatches   52;   Indels   6;   Gaps   2;

Qy   40   RAKEKRNRLSLLVQKPEFHEDTRSS-----RSGHLAKETRVSPEEAVKWGESFDKLLSHR 94
        : || |||:: | :|:|| :::: || |||:| || |:
Db   15   RAKEFKTRLGIFLHKSELGCDTGSTGKFEWGSKH-SKENRNFSEDVLGWRESFDLLLSSK 73

Qy   95   DGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLD 154
        :|: ||||||||||||| | |: |||| : : :|:|| ||||||:|
Db   74   NGVAAFHAFLKTEFSEENLEFWLACEEFKKIRSATKLASRAHQIFEEFICSEAPKEVNID 133

Qy   155   FHTKEVITNSITQPTLHSFDAAQSRVVQLMEQDSYTRFLKSDIYLDL 201
        |:|: :: :||| |||||
Db   134   HETRELTRMNLQTATATCFDAAQGKTRTLMEKDSYPRFLKSPAYRDL 180

RESULT 13
US-08-754-108-2
; Sequence 2, Application US/08754108
; Patent No. 5886149
; GENERAL INFORMATION:
;   APPLICANT: Buckbinder, Leonard
;   APPLICANT: Talbott, Randy
;   APPLICANT: Seizinger, Bernd R.
;   APPLICANT: Kley, Nikolai
;   TITLE OF INVENTION: No. 58861149el p53 Response Genes
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Bristol-Myers Squibb Company
;   STREET: 3005 First Avenue
;   CITY: Seattle
;   STATE: Washington
;   COUNTRY: USA
;   ZIP: 98121
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0,   Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/754,108
;   FILING DATE:
;   CLASSIFICATION: 530
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/274,318
;   FILING DATE: 11-JUL-1994
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Bogden, James M.
;   REGISTRATION NUMBER: 32,962

;   REFERENCE/DOCKET NUMBER: ON0127-
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 206-727-3688
;   TELEFAX: 206-727-3601
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 202 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-754-108-2

Query Match   31.2%;   Score 380;   DB 2;   Length 202;
Best Local Similarity   48.5%;   Pred. No. 6.9e-31;
Matches   81;   Conservative   28;   Mismatches   52;   Indels   6;   Gaps   2;

Qy   40   RAKEKRNRLSLLVQKPEFHEDTRSS-----RSGHLAKETRVSPEEAVKWGESFDKLLSHR 94
        : || |||:: | :|:|| :::: || |||:| || |:
Db   15   RAKEFKTRLGIFLHKSELGCDTGSTGKSEWGSKH-SKENRNFSEDVLGWRESFDLLLSSK 73

Qy   95   DGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLD 154
        :|: ||||||||||||| | |: |||| : : :|:|| ||||||:|
Db   74   NGVAAFHAFLKTEFSEENLEFWLACEEFKKIRSATKLASRAHQIFEEFICSEAPKEVNID 133

Qy   155   FHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDL 201
        |:|: :: :||| |||||
Db   134   HETRELTRMNLQTATATCFDAAQGKTRTLMEKDSYPRFLKSPAYRDL 180

RESULT 14
US-08-870-815-4
; Sequence 4, Application US/08870815
; Patent No. 6020142
; GENERAL INFORMATION:
;   APPLICANT: Levinson, Douglas A.
;   APPLICANT: Gimeno, Carlos J.
;   TITLE OF INVENTION: RATH GENES AND POLYPEPTIDES AND METHODS
;   TITLE OF INVENTION: FOR THE TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
;   NUMBER OF SEQUENCES: 15
;   CORRESPONDENCE ADDRESS:
;   ADDRESSE: Pennie & Edmonds LLP
;   STREET: 1155 Avenue of the Americas
;   CITY: New York
;   STATE: New York
;   COUNTRY: USA
;   ZIP: 10036-2711
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0,   Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/870,815
;   FILING DATE: 06-JUN-1997
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Laura A. Coruzzi
;   REGISTRATION NUMBER: 30,742
;   REFERENCE/DOCKET NUMBER: 7853-074
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 790-9090
;   TELEFAX: (212) 896-8864/9741
;   TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 202 amino acids
;   TYPE: amino acid
;   STRANDEDNESS:
;   TOPOLOGY: unknown
;   MOLECULE TYPE: protein
US-08-870-815-4

Query Match   31.2%;   Score 380;   DB 3;   Length 202;

Best Local Similarity 48.5%; Pred. No. 6.9e-31;
Matches   81; Conservative   28; Mismatches   52; Indels   6; Gaps   2;

Qy   40 RAKEKRNRLSLLVQKPEFHEDTRSS-----RSGHLAKETRVSPEEAVKWGESFDKLLSHR 94
            |||| :|| :|| :| | | : :| : || |||| :: ||| ||| |||||| :
Db   15 RAKEFKTRLGIFLHKSELGCDTGSTGKFEWGSKH-SKENRNFSEDVLGWRESFDLLLSSK 73

Qy   95 DGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLD 154
            :: | ||||||||||| |||||::||| : :: :|:| ||:|||||||:|||||||
Db   74 NGVAAFHAFLKTEFSEENLEFWLACEEFKKIRSATKLASRAHQIFEEPICSEAPKEVNID 133

Qy   155 FHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDL 201
            |:|: : :: |:|| |:|||| |||||||||||| |||| || ||
Db   134 HETRELTRMNLQTATATCFDAAQGKTRTLMEKDSYPRFLKSPAYRDL 180

RESULT 15
US-08-463-074B-2
; Sequence 2, Application US/08463074B
; Patent No. 6020155
; GENERAL INFORMATION:
;   APPLICANT: Smith, Kendall A. & Beadling, Carol
;   TITLE OF INVENTION: Nucleic Acids Encoding CR1 Fusion Protein, Vector an
;   NUMBER OF SEQUENCES: 35
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
;   ADDRESSEE: (B) STREET: 444 South Flower St. - Suite 1900
;   CITY: Los Angeles
;   STATE: California
;   COUNTRY: USA
;   ZIP: 90071
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0,
;   SOFTWARE: Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/463,074B
;   FILING DATE: 5-JUN-1995
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/104,736
;   FILING DATE: 10-AUG-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/796,066
;   FILING DATE: 20-NOV-91
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Viviana Amzel, Ph.D.
;   REGISTRATION NUMBER: 30,930
;   REFERENCE/DOCKET NUMBER: P66 38143 (DART-020)
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (213) 622-7700
;   TELEFAX: (213) 489-4210
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 202 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-463-074B-2

Query Match   31.2%; Score 380; DB 3; Length 202;
Best Local Similarity 48.5%; Pred. No. 6.9e-31;
Matches   81; Conservative   28; Mismatches   52; Indels   6; Gaps   2;

Qy   40 RAKEKRNRLSLLVQKPEFHEDTRSS-----RSGHLAKETRVSPEEAVKWGESFDKLLSHR 94
            |||| :|| :|| :| | | : :| : || |||| :: ||| ||| |||||| :
Db   15 RAKEFKTRLGIFLHKSELGCDTGSTGKFEWGSKH-SKENRNFSEDVLGWRESFDLLLSSK 73

Qy   95 DGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLD 154
            :: | ||||||||||| |||||::||| : :: :|:| ||:|||||||:|||||||
Db   74 NGVAAFHAFLKTEFSEENLEFWLACEEFKKIRSATKLASRAHQIFEEPICSEAPKEVNID 133

Qy   155 FHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDL 201
            |:|: : :: |:|| |:|||| |||||||||||| |||| || ||
Db   134 HETRELTRMNLQTATATCFDAAQGKTRTLMEKDSYPRFLKSPAYRDL 180

Search completed: August 20, 2004, 16:48:32
Job time : 19.5 secs

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:        August 20, 2004, 16:47:08 ; Search time 239 Seconds
                                (without alignments)
                                308.996 Million cell updates/sec

Title:            US-09-894-749-2
Perfect score:    1218
Sequence:         1 METTLLFFSQINMCESKEKT.........SRSFTCNEFQDVQSDVAIWL 235
Scoring table:    BLOSUM62
                  Gapop 10.0 , Gapext 0.5
Searched:         1295152 seqs, 314255058 residues
Total number of hits satisfying chosen parameters:        1295152
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :        Published Applications_AA:*
    1: /cgn2_6/ptcdata/2/pubpaa/US07_PUBCOMB.pep:*
    2: /cgn2_6/ptcdata/2/pubpaa/PCT_NEW_PUB.pep:*
    3: /cgn2_6/ptcdata/2/pubpaa/US06_NEW_PUB.pep:*
    4: /cgn2_6/ptcdata/2/pubpaa/US06_PUBCOMB.pep:*
    5: /cgn2_6/ptcdata/2/pubpaa/US07_NEW_PUB.pep:*
    6: /cgn2_6/ptcdata/2/pubpaa/PCTUS_PUBCOMB.pep:*
    7: /cgn2_6/ptcdata/2/pubpaa/US08_NEW_PUB.pep:*
    8: /cgn2_6/ptcdata/2/pubpaa/US08_PUBCOMB.pep:*
    9: /cgn2_6/ptcdata/2/pubpaa/US09A_PUBCOMB.pep:*
   10: /cgn2_6/ptcdata/2/pubpaa/US09B_PUBCOMB.pep:*
   11: /cgn2_6/ptcdata/2/pubpaa/US09C_PUBCOMB.pep:*
   12: /cgn2_6/ptcdata/2/pubpaa/US09_NEW_PUB.pep:*
   13: /cgn2_6/ptcdata/2/pubpaa/US10A_PUBCOMB.pep:*
   14: /cgn2_6/ptcdata/2/pubpaa/US10B_PUBCOMB.pep:*
   15: /cgn2_6/ptcdata/2/pubpaa/US10C_PUBCOMB.pep:*
   16: /cgn2_6/ptcdata/2/pubpaa/US10_NEW_PUB.pep:*
   17: /cgn2_6/ptcdata/2/pubpaa/US60_NEW_PUB.pep:*
   18: /cgn2_6/ptcdata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1218	100.0	235	9	US-09-894-749-2	Sequence 2, Appli
2	1218	100.0	235	12	US-10-258-371B-20	Sequence 20, App
3	1176	96.6	227	9	US-09-867-550-848	Sequence 848, App
4	1032	84.7	235	9	US-09-894-749-4	Sequence 4, Appli
5	422.5	34.7	211	9	US-09-206-639-4	Sequence 4, Appli
6	422.5	34.7	211	12	US-10-258-371B-24	Sequence 24, Appl
7	422.5	34.7	211	16	US-10-405-765A-493	Sequence 493, App
8	422.5	34.7	220	9	US-09-925-300-1507	Sequence 1507, Ap
9	421.5	34.6	181	9	US-09-206-639-1	Sequence 1, Appli
10	421.5	34.6	181	12	US-10-258-371B-22	Sequence 22, Appl
11	421.5	34.6	190	14	US-10-411-224-197	Sequence 197, App
12	421.5	34.6	190	14	US-10-047-021-197	Sequence 197, App
13	413	33.9	81	12	US-10-258-371B-12	Sequence 12, Appli
14	410.5	33.7	181	14	US-10-398-953-1	Sequence 1, Appli
15	409	33.6	930	14	US-10-113-794A-1	Sequence 1, Appli
16	408.5	33.5	205	9	US-09-206-639-5	Sequence 5, Appli
17	408.5	33.5	205	10	US-09-939-209A-2	Sequence 2, Appli
18	408.5	33.5	205	12	US-10-258-371B-21	Sequence 21, Appl
19	404	33.2	519	12	US-10-258-371B-28	Sequence 28, Appl
20	404	33.2	519	14	US-10-113-794A-2	Sequence 2, Appli
21	404	33.2	519	15	US-10-428-487-14	Sequence 14, Appl
22	404	33.2	591	15	US-10-108-260A-3970	Sequence 3970, Ap
23	402.5	33.0	180	12	US-10-258-371B-27	Sequence 27, Appl
24	402.5	33.0	180	12	US-10-275-555-2	Sequence 2, Appli
25	401	32.9	776	12	US-10-087-192-1728	Sequence 1728, Ap
26	400	32.8	284	15	US-10-094-749-1650	Sequence 1650, Ap
27	396.5	32.6	923	12	US-10-114-270-152	Sequence 152, App
28	380	31.2	202	9	US-09-873-438-4	Sequence 4, Appli
29	380	31.2	202	14	US-10-171-311-202	Sequence 202, App
30	380	31.2	204	12	US-10-258-371B-23	Sequence 23, Appl
31	370	30.4	128	15	US-10-352-843-1	Sequence 1, Appli
32	370	30.4	128	15	US-10-352-843-3	Sequence 3, Appli
33	370	30.4	167	10	US-09-942-055A-1	Sequence 1, Appli
34	361	29.6	69	12	US-10-258-371B-13	Sequence 13, Appli
35	361	29.6	201	9	US-09-873-438-2	Sequence 2, Appli
36	353	29.0	196	9	US-09-806-639-3	Sequence 3, Appli
37	353	29.0	196	9	US-09-736-457-339	Sequence 339, App
38	353	29.0	196	9	US-09-902-941-339	Sequence 339, App
39	353	29.0	196	9	US-09-849-626-339	Sequence 339, App
40	353	29.0	196	10	US-09-476-300-339	Sequence 339, App
41	353	29.0	196	14	US-10-283-017-339	Sequence 339, App
42	353	29.0	196	14	US-10-017-754-339	Sequence 339, App
43	353	29.0	196	14	US-10-113-872-339	Sequence 339, App
44	353	29.0	196	14	US-10-247-671-176	Sequence 176, App
45	353	29.0	217	9	US-09-925-301-1292	Sequence 1292, Ap

ALIGNMENTS

```
RESULT 1
US-09-894-749-2
; Sequence 2, Application US/0989749
; Patent No. US20020081683A1
; GENERAL INFORMATION:
;  APPLICANT: Hodge, Martin R.
;  APPLICANT: Yowe, David
;  TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
;  FILE REFERENCE: 5800-19, 035800/174680
;  CURRENT APPLICATION NUMBER: US/09/894,749
;  CURRENT FILING DATE: 2001-06-27
;  PRIOR APPLICATION NUMBER: 09/244,314
;  PRIOR FILING DATE: 1999-02-04
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;  LENGTH: 235
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-09-894-749-2

Query Match             100.0%;  Score 1218;  DB 9;  Length 235;
Best Local Similarity   100.0%;  Pred. No. 6.7e-110;
Matches   235;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNLSLLVQKPEFHED 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNLSLLVQKPEFHED 60

Qy     61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTFSEENIEFWIACE 120

Qy    121 DFKKSKGPQQIHLKAKAIYEKPIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 DFKKSKGPQQIHLKAKAIYEKPIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy    181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
```

```
Db  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 2
US-10-258-371B-20
; Sequence 20, Application US/10258371B
; Publication No. US20040067903A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS-GAGNON, Alison
; APPLICANT: MURRAY, David L
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING,
; TITLE OF INVENTION: RGS18, AND USES THEREOF
; FILE REFERENCE: A3656 US PCT
; CURRENT APPLICATION NUMBER: US/10/258,371B
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: GB01883.334
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US60/200,786
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-371B-20

Query Match   100.0%; Score 1218; DB 12; Length 235;
Best Local Similarity 100.0%; Pred. No. 6.7e-110;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
```

RESULT 3
US-09-867-550-848
; Sequence 848, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US2002008220206A1el Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastcSEQ for Windows Version 4.0
; SEQ ID NO 848
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-848

Query Match   96.6%; Score 1176; DB 9; Length 227;
Best Local Similarity 100.0%; Pred. No. 7.7e-106;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDV 227
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDV 227
```

RESULT 4
US-09-894-749-4
; Sequence 4, Application US/09894749
; Patent No. US20020081683A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; APPLICANT: Yowe, David
; TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
; FILE REFERENCE: 5800-19, 035800/174680
; CURRENT APPLICATION NUMBER: US/09/894,749
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/244,314
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-894-749-4

Query Match   84.7%; Score 1032; DB 9; Length 235;
Best Local Similarity 83.8%; Pred. No. 7.7e-92;
Matches 197; Conservative 19; Mismatches 19; Indels 0; Gaps 0;

```
Qy    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
          |:|| |||||:|||||||||||||:|||||||| ||||||||||||||||| |:||:|
Db    1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60

Qy   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
          |::|| :||||||||||||||||:||||||||||||:|:|||||||||||||||||:||
Db   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120

Qy  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
          ||||   ||||:||||||||||||.|||||||:|||||||| :|||:|||||||||:||||
Db  121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180

Qy  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
          |||||:|||:||||| |||:|:|:||||||||||||||||| |:|||||:||||||
Db  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
```

RESULT 5
US-09-206-639-4
; Sequence 4, Application US/09206639
; Patent No. US20020034777A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: A NOVEL REGULATOR OF CELL SIGNALLING
; NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/206,639
; FILING DATE: 07-Dec-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/748,483
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0157 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 211 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 292037
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-206-639-4

Query Match　　34.7%;　Score 422.5; DB 9;　Length 211;
Best Local Similarity　42.8%;　Pred. No. 1e-32;
Matches　89;　Conservative　37;　Mismatches　73;　Indels　9;　Gaps　3;

Qy　　6　LFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTR 62
　　　　　:|: :||::::　||||　　|||　||　:　||||:：||　|
Db　　5　MFLAVQHDCRPMDKS----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSSTPGKPK 59

Qy　　63　SSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDF 122
　　　　　: :|| :|||||||　|::||||||:　||：|||||:||||||
Db　　60　TGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACEDF 119

Qy　　123　KKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQ 182
　　　　　||||::||:　　||:||::||:||||::||:：|| :::||| ||||
Db　　120　KKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRVYS 179

Qy　　183　LMEQDSYTRFLKSDIYLDLMEGRPQRPT 210
　　　　　||| :||:|:||||::||:
Db　　180　LMENNSYPRFLESEFYQDLCK-KPQITT 206

RESULT 6
US-10-258-371B-24
; Sequence 24, Application US/10258371B
; Publication No. US2004006790A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS-GAGNON, Alison
; APPLICANT: MURRAY, David L
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING,
; TITLE OF INVENTION: RGS18, AND USES THEREOF
; FILE REFERENCE: A3656 US PCT
; CURRENT APPLICATION NUMBER: US/10/258,371B
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: GB001883.334

; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US60/200,786
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-371B-24

Query Match　　34.7%;　Score 422.5; DB 12;　Length 211;
Best Local Similarity　42.8%;　Pred. No. 1e-32;
Matches　89;　Conservative　37;　Mismatches　73;　Indels　9;　Gaps　3;

Qy　　6　LFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTR 62
　　　　　:|: :||::::　||||　　|||　||　:　||||:：||　|
Db　　5　MFLAVQHDCRPMDKS----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSSTPGKPK 59

Qy　　63　SSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDF 122
　　　　　: :|| :|||||||　|::||||||:　||：|||||:||||||
Db　　60　TGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACEDF 119

Qy　　123　KKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQ 182
　　　　　||||::||:　　||:||::||:||||::||:：|| :::||| ||||
Db　　120　KKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRVYS 179

Qy　　183　LMEQDSYTRFLKSDIYLDLMEGRPQRPT 210
　　　　　||| :||:|:||||::||:
Db　　180　LMENNSYPRFLESEFYQDLCK-KPQITT 206

RESULT 7
US-10-408-765A-493
; Sequence 493, Application US/10408765A
; Publication No. US2004010187A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Sounitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 493
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-493

Query Match　　34.7%;　Score 422.5; DB 16;　Length 211;
Best Local Similarity　42.8%;　Pred. No. 1e-32;
Matches　89;　Conservative　37;　Mismatches　73;　Indels　9;　Gaps　3;

Qy　　6　LFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTR 62
　　　　　:|: :||::::　||||　　|||　||　:　||||:：||　|
Db　　5　MFLAVQHDCRPMDKS----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSSTPGKPK 59

Qy　　63　SSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDF 122
　　　　　: :|| :|||||||　|::||||||:　||：|||||:||||||
Db　　60　TGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACEDF 119

Qy　　123　KKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQ 182
　　　　　||||::||:　　||:||::||:||||::||:：|| :::||| ||||
Db　　120　KKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRVYS 179

Qy　　183　LMEQDSYTRFLKSDIYLDLMEGRPQRPT 210
　　　　　||| :||:|:||||::||:
Db　　180　LMENNSYPRFLESEFYQDLCK-KPQITT 206

```
Qy 183 LMEQDSYTRFLKSDIYLDLMEGRPQRPT 210
          ||||:||||:||:||: |||||
Db 180 LMENNSYPRFLESEFYQDLCK-KPQITT 206

RESULT 8
US-09-925-300-1507
; Sequence 1507, Application US/09925300
; Patent No. US2002015168lA1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1507
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1507

Query Match         34.7%; Score 422.5; DB 9; Length 220;
Best Local Similarity 42.8%; Pred. No. 1.1e-32;
Matches 89; Conservative 37; Mismatches 73; Indels 9; Gaps 3;

Qy   6 LFFSQIMCESKEKTFFKLIHGSGKEETSKEAKIR---AKEKRNRLSLLVQKPEFHEDTR 62
          :||:|                            ||:|:|||||:||:||:
Db  14 MFLAVQHDCRPMDKS-----AGSGHKSEKREKMKRTLLKDWKTRLSYFLQNSSTPGKPK 68

Qy  63 SSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDF 122
          :|||:||||:|||||||||||||||||:|||||    |||:|||:|||:|||||||||
Db  69 TGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACEDF 128

Qy 123 KKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQ 182
          :||:||||| |: :|:||||| |||||
Db 129 KKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRVYS 188

Qy 183 LMEQDSYTRFLKSDIYLDLMEGRPQRPT 210
          ||||:||||:||:||: |||||
Db 189 LMENNSYPRFLESEFYQDLCK-KPQITT 215

RESULT 9
US-09-206-639-1
; Sequence 1, Application US/09206639
; Patent No. US2002034777A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
;            Goli, Surya K.
; TITLE OF INVENTION: A NOVEL REGULATOR OF CELL SIGNALLING
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/206,639
; FILING DATE: 07-Dec-1998
```

```
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/748,483
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36.749
; REFERENCE/DOCKET NUMBER: PF-0157 US
; TELECOMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: 57362
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-206-639-1

Query Match         34.6%; Score 421.5; DB 9; Length 181;
Best Local Similarity 49.1%; Pred. No. 1e-32;
Matches 82; Conservative 37; Mismatches 45; Indels 3; Gaps 1;

Qy  40 RAKEKRNRLSLLVQKPEFHED---TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG 96
         ||||  ||: :|:||||:         ::|||||||||||:||||||||:||||||||
Db  15 RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG 74

Qy  97 LEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
         |||| |:                 :||:||||:||
Db  75 LASFKSFLKSBFSEENLEFWIACEDYKKIKSPAKMAEKAKQIYEFIQTEAPKEVNIDHF 134

Qy 157 TKEVITNSITQPTLHSFDAAQSRVQLMEQDSYTRFLKSDIYLDLME 203
         |::|:    :||:|||||||:||| :|::|:: :|:
Db 135 TKDITMKNIVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK 181

RESULT 10
US-10-258-371B-22
; Sequence 22, Application US/10258371B
; Publication No. US20040067903A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS-GAGNON, Alison
; APPLICANT: MURRAY, David L
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING,
; TITLE OF INVENTION: RGS18, AND USES THEREOF
; FILE REFERENCE: A3656 US PCT
; CURRENT APPLICATION NUMBER: US/10/258,371B
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: GB001883.334
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US60/200,786
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-371B-22

Query Match         34.6%; Score 421.5; DB 12; Length 181;
Best Local Similarity 49.1%; Pred. No. 1e-32;
Matches 82; Conservative 37; Mismatches 45; Indels 3; Gaps 1;

Qy  40 RAKEKRNRLSLLVQKPEFHED---TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG 96
         ||||  ||: :|:||||:         ::|||||||||||:||||||||:||||||||
Db  15 RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG 74
```

```
Qy   97  LEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH  156
         |:| ||:|||||||||||||||| :: ||| ::|||||||:|||:|:|
Db   75  LASFKSFLKSEFSEENLEFWIACEDYKKKIKSPAKMAEKAKQIYEEFIQTEAPKEVNIDHF  134

Qy  157  TKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLME  203
         ||:|: |||||||:||||||||||||||||||||:|||
Db  135  TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK  181
```

RESULT 11
US-10-411-224-197
; Sequence 197, Application US/10411224
; Publication No. US20030166906A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 50 Human Secreted Proteins
; FILE REFERENCE: P2016P1
; CURRENT APPLICATION NUMBER: US/10/411,224
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 09/722,329
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/262,109
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: 60/057,626
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/057,663
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/057,669
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/058,667
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,973
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,666
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,974
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 197
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-224-197

Query Match   34.6%; Score 421.5; DB 14; Length 190;
Best Local Similarity   49.1%; Pred. No. 1.1e-32;
Matches   82; Conservative 37; Mismatches 45; Indels   3; Gaps   1;

Qy   40 RAKEKRNRLSLLVQKPEFHED--TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG 96
         ||||| ||:|:|||:|:| : : :|:|::| |||| :| :|
Db   24 RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG 83

Qy   97 LEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
         |:| ||:|||||||||||||||| :: ||| ::|||||||:|||:|:|
Db   84 LASFKSFLKSEFSEENLEFWIACEDYKKKIKSPAKMAEKAKQIYEEFIQTEAPKEVNIDHF 143

Qy 157 TKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLME 203
         ||:|: |||||||:||||||||||||||||||||:|||
Db 144 TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK 190
```

RESULT 12
US-10-047-021-197
; Sequence 197, Application US/10047021
; Publication No. US20040002591A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 50 Human Secreted Proteins
; FILE REFERENCE: P2016P2
; CURRENT APPLICATION NUMBER: US/10/047,021
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/262,066
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 09/722,329
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/262,109
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: PCT/US98/18360
; PRIOR FILING DATE: 1998-09-03
; PRIOR APPLICATION NUMBER: US 60/057,626
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/057,663
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/057,669
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/058,667
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,974
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,973
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,666
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/090,112
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 197
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-047-021-197

Query Match   34.6%; Score 421.5; DB 15; Length 190;
Best Local Similarity   49.1%; Pred. No. 1.1e-32;
Matches   82; Conservative 37; Mismatches 45; Indels   3; Gaps   1;

Qy   40 RAKEKRNRLSLLVQKPEFHED--TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG 96
         ||||| ||:|:|||:|:| : : :|:|::| |||| :| :|
Db   24 RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG 83

Qy   97 LEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
         |:| ||:|||||||||||||||| :: ||| ::|||||||:|||:|:|
Db   84 LASFKSFLKSEFSEENLEFWIACEDYKKKIKSPAKMAEKAKQIYEEFIQTEAPKEVNIDHF 143

Qy 157 TKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLME 203
         ||:|: |||||||:||||||||||||||||||||:|||
Db 144 TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK 190
```

RESULT 13
US-10-258-371B-12
; Sequence 12, Application US/10258371B
; Publication No. US20040067903A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS-GAGNON, Alison
; APPLICANT: MURRAY, David L
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING,
; TITLE OF INVENTION: RGS18, AND USES THEREOF
; FILE REFERENCE: A3656 US PCT
; CURRENT APPLICATION NUMBER: US/10/258,371B
; CURRENT FILING DATE:  2003-06-04
; PRIOR APPLICATION NUMBER: GB001883,334
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US60/200,786
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-371B-12
```

```
Matches  '82;  Conservative   37;  Mismatches   60;  Indels    2;  Gaps    1;

Qy     31 EETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTRSSRGHLAKETRVSPEEAVKWGESFD  88
          :||::  ||::.:|  ||| .:|::.  ::  ||:|  ||||   |:|:.:  |||||
Db    748 DEASRKRKSKNIAKDMKNKLAIFRRRNESPGAQPASKTDKTTKSFKPTSEEALKWSESLE 807

Qy     89 KLLSHRDGLEAFTRPLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAP 148
          :||  |:.:|||::||||||:|:||||||||||||||||||||||||||||:||:||
Db    808 KLLLHKYGLEVFQAFLRTEFSEENLEFWLACEDFKKVKSQSKMAAKAKKIFAEFIAIQAC 867

Qy    149 KEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQR 208
          ||||||||||  :||:  ||| .:: |||||||||||||:: |||||||||||:  :
Db    868 KEVNLDSYTREHTKENLQSITRGCFDLAQKRIFGLMEKDSYPRFLRSDLYLDLLINQKKMS 927

Qy    209 P 209
          P
Db    928 P 928

Search completed: August 20, 2004, 17:06:39
Job time : 241 secs
```

```
Query Match             33.9%;      Score 413;  DB 12;  Length 81;
Best Local Similarity   97.5%;      Pred. No. 2.3e-32;
Matches   78;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    110 EENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPT 169
          |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 EENLEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPT 61

Qy    170 LHSFDAAQSRVYQLMEQDSY 189
          |||||||||||||||||||||
Db     62 LHSFDAAQSRVYQLMENDSY 81

RESULT 14
US-10-398-953-1
; Sequence 1, Application US/10398953
; Publication No. US2003018630A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: No. US2003018630A1e1 disease-associated gene and use thereof
; FILE REFERENCE: 2801 USOP
; CURRENT APPLICATION NUMBER: US/10/398,953
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: JP 2000-319912
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: JP 2000-350183
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
;   LENGTH: 181
;   TYPE: PRT
;   ORGANISM: Rat
US-10-398-953-1
```

```
Query Match             33.7%;      Score 410.5;  DB 14;  Length 181;
Best Local Similarity   48.2%;      Pred. No. 1.2e-31;
Matches   80;  Conservative   35;  Mismatches   48;  Indels    3;  Gaps    1;

Qy     40 RAKEKRNRLSLLVQKPEFHED--TRSSRGHLAKETRVSPEEAVKWGESFDKLLSHRDG 96
          |||| :::| :|:||:: ||| .  : .:|.:::| ||||| |:|:.:||  ||||
Db     15 RAKEIKIKLGILLQKPDSAVDLVIPYNEKPEKPAKAHKPSLEEVLQWRQSLDKLLQSNYG 74

Qy     97 LEAFTRPLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
          ||| :|| |:|:||||||||  ||||||:  |:|||||||||| |||||||||||||:
Db     75 FASFKSFLKSEFSEENLEFWVACENYKKIKSPIKMABKAKQIYEEFIQTEAPKEVNIDHF 134

Qy    157 TKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLM 202
          ||| :|| ||| |||:|||||||||||||  ||||||||||||:|
Db    135 TKDITMKNLVEPSPHSFDLAQKRIYALMEKDSLPRFVRSEFYKELI 180

RESULT 15
US-10-113-794A-1
; Sequence 1, Application US/10113794A
; Publication No. US2003002202A1
; GENERAL INFORMATION:
; APPLICANT: Flanagan et al.
; TITLE OF INVENTION: B EPHRIN REGULATION OF G-PROTEIN COUPLED
; TITLE OF INVENTION: CHEMOATTRACTION
; FILE REFERENCE: 2535/106
; CURRENT APPLICATION NUMBER: US/10/113,794A
; CURRENT FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 930
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-113-794A-1

Query Match             33.6%;      Score 409;  DB 14;  Length 930;
Best Local Similarity   45.3%;      Pred. No. 1.6e-30;
```

```
              GenCore version 5.1.6
        Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:    August 20, 2004, 16:14:26 ; Search time 25 Seconds
                              (without alignments)
                    904.200 Million cell updates/sec

Title:         US-09-894-749-2
Perfect score: 1218
Sequence:      1 METTLLFFSQINMCESKEKT.........SRSFTCNEFQDVQSDVAIWL 235
Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters:        283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    PIR_78:*
              1: pir1:*
              2: pir2:*
              3: pir3:*
              4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

             %
Result       Query
No.   Score  Match  Length  DB  ID        Description
----  -----  -----  ------  --  --------  -----------
1     422.5  34.7   211     2   I53020    G-0/G-1 switch reg
2     408.5  33.5   205     2   S78221    G-protein signalin
3     404    33.2   519     2   S78089    G-protein signalin
4     400.5  32.9   181     2   JC7228    G-protein signalin
5     346    28.4   196     2   S43436    B cell activation
6     315.5  25.9   181     2   T21035    hypothetical prote
7     313.5  25.7   251     2   S43576    C05B5.7 protein (c
8     313.5  25.7   254     2   G88571    protein C05B5.7 [i
9     313    25.7   169     2   T21034    hypothetical prote
10    312    25.6   244     2   T13580    RGS10 protein - hu
11    282.5  23.2   173     2   S71812    hypothetical prote
12    278.5  22.9   533     2   T31002    G-protein signalin
13    243.5  20.0   544     2   JC5503    G-protein signalin
14    225    18.5   1387    2   JC5502    hypothetical prote
15    220    18.1   473     2   T19337    hypothetical prote
16    216    17.7   558     2   T21468    hypothetical prote
17    215.5  17.7   234     2   T26672    hypothetical prote
18    208    17.1   838     2   T08423    Axin homolog Axil
19    194.5  16.0   832     2   T08422    negative regulator
20    178.5  14.7   270     2   T22213    hypothetical prote
21    178.5  14.7   284     2   T15700    hypothetical prote
22    158.5  13.0   274     2   T22214    hypothetical prote
23    149    12.2   235     2   F89472    protein ZC53.7 [im
24    141    11.6   719     2   S60771    developmental regu
25    135    11.1   303     2   T29513    hypothetical prote
26    116    9.5    229     2   A89473    protein F52D2.2 [i
27    114    9.4    1010    2   T41077    hypothetical struc
28    113.5  9.3    309     2   S61665    probable membrane
29    102.5  8.4    688     1   JC1469    beta-adrenergic-re
30    101.5  8.3    688     1   A39336    beta-adrenergic-re
31    98     8.0    1092    2   T12520    hypothetical prote
32    97     8.0    423     2   T15335    hypothetical prote
33    97     8.0    913     2   T31497    hypothetical prote
34    96.5   7.9    950     2   F86286    probable membrane
35    95.5   7.8    1246    2   S60954    hypothetical prote
36    95     7.8    1187    2   T18355    hypothetical prote
37    94     7.7    997     1   I54390    phosphoinositide-s
38    94     7.7    1319    2   A28313    glued protein - fr
39    93.5   7.7    1514    2   S70099    hypothetical prote
40    93     7.6    258     2   T00447    hypothetical prote
41    92.5   7.6    535     2   T37189    hypothetical prote
42    92.5   7.6    537     2   E96681    protein F1E22.4 [i
43    91.5   7.5    393     2   H83747    two-component sens
44    91.5   7.5    480     2   S72668    oviduct-specific p
45    91.5   7.5    3418    1   G02334    breast cancer tumo
```

                              ALIGNMENTS

```
RESULT 1
I53020
G-0/G-1 switch regulatory protein 8 - human
N;Alternate names: helix-loop-helix phosphoprotein; regulator of G-protein signaling 2, ;
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C;Accession: I53020; I65984
R;Sidorovski, D.P.; Hexiner, S.P.; Forsdyke, D.R.
DNA Cell Biol. 13, 125-147, 1994
A;Title: A human gene encoding a putative basic helix-loop-helix phosphoprotein whose mRN
A;Reference number: I53020; MUID:94235118; PMID:8179820
A;Accession: I53020
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <RES>
A;Cross-references: GB:L13391; NID:g292036; NID:g292036.1; PIDN:AAA20680.1; PID:g292037
A;Accession: I65984
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-211 <RE2>
A;Cross-references: GB:L13463; NID:g292054; NID:g292054.1; PIDN:AAC37587.1; PID:g292055
C;Genetics:
A;Gene: GDB:RGS2; GOS8
A;Cross-references: GDB:355647; OMIM:600861
A;Map position: 1q31-1q31
A;Introns: 37/2; 71/2; 92/1; 147/3
C;Superfamily: B-cell activation protein BL34
C;Keywords: phosphoprotein

Query Match           34.7%;   Score 422.5;  DB 2;  Length 211;
Best Local Similarity 42.8%;   Pred. No. 3.3e-25;
Matches  89; Conservative  37; Mismatches  73; Indels   9; Gaps   3;

Qy      6 LFFSQINMCESKEKTFFKLIHGSGKSETSKEAKIR---AKEKRNRLSLLVQKPEFHEDTR 62
            |||:  :|: . |||||  ||||||  | |: :|
Db      5 MFLAVQHDCRPMDKS-----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSSTPGKPK 59

Qy     63 SSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDF 122
           ||:||:|:||  ||: ||||  ||| ||||::|  ||:|||||||| ||||| |||||||
Db     60 TGKKSKQQAFIKPSPEAQLWSEAFDLLASKYGLAAFRAFIKSEFCEENIEFWLACEDF 119

Qy    123 KKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQ 182
           |||||||||:|:|:|:||||||  |:|||||:||||||||| ::|||||||:||||||
Db    120 KKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQKTLIAQNIQRATSGCFTTAQXRVYS 179

Qy    183 LMEQDSYTRFLKSDIYLDLMEGRPQRPT 210
           |||||||||||:::||| || :|| :|:|
Db    180 LMENNSYPRFLESEFYQDLCK-KPQITT 206

RESULT 2
S78221
```

G-protein signaling regulator RGP4 - human
N;Alternate names: G-protein signaling regulator RGS4
C;Species: Homo sapiens (man)
C;Date: 29-Jan-1998 #sequence_revision 13-Feb-1998 #text_change 31-Jan-2000
R;Druey, K.
C;Accession: S78221; S68509
submitted to the EMBL Data Library, May 1995
A;Reference number: S78089
A;Accession: S78221
A;Molecule type: mRNA
A;Residues: 1-205 <DRU>
A;Cross-references: EMBL:U27768; NID:g1216372; PIDN:AAC50395.1; PID:g1216373
R;Druey, K.M.; Blumer, K.J.; Kang, V.H.; Kehrl, J.H.
Nature 379, 742-746, 1996
A;Title: Inhibition of G-protein-mediated MAP kinase activation by a new mammalian gene
A;Reference number: A58012; MUID:96178495; PMID:8602223
A;Accession: S68509
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-35,'S',37-39,'S',41-46,'T',48-67,'N',69-76,'L',78-177,'T',179-197,'T',199-
A;Cross-references: EMBL:U27768
A;Experimental source: brain
C;Superfamily: B-cell activation protein BL34

```
Query Match          33.5%;  Score 408.5;  DB 2;  Length 205;
Best Local Similarity 49.1%;  Pred. No. 3.7e-24;
Matches  82;  Conservative  28;  Mismatches  48;  Indels  9;  Gaps  2;

Qy  41 AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKE----TRVSPEEAVKWGESFDKLLSHRD 95
        ||| |:            :||            ||||
Db  16 AKDMKHRLGFLLQK---SDSCEHNSSHNKDKVVICQRVSQEEVKWAESLENIISHEC 71

Qy  96 GLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDF 155
        || |:: ||||:          ||     :|  ||||||
Db  72 GLAAFKAFLKSEYSEENIDFWISCEEYKKIKSPSKLSPKAKKIYNEFISVQATKEVNLDS 131

Qy 156 HTKEVITNSITQPTLHSFPDAAQSRVQLMEQDSYTRFLKSDIYLDLM 202
        |||  :: |||  :|    |||||| |||||
Db 132 CTREETSRNMLEPTITCFDEAQKKIFNLMEKDSYRRFLKSRFYLDLV 178
```

RESULT 3
S78089
G-protein signaling regulator RGS3 - human
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 24-Nov-1999
R;Druey, K.
C;Accession: S78089; S68436
submitted to the EMBL Data Library, May 1995
A;Reference number: S78089
A;Accession: S78089
A;Molecule type: mRNA
A;Residues: 1-519 <DRU>
A;Cross-references: EMBL:U27655; NID:g1216368; PID:g1216369
R;Druey, K.M.; Blumer, K.J.; Kang, V.H.; Kehrl, J.H.
Nature 379, 742-746, 1996
A;Title: Inhibition of G-protein-mediated MAP kinase activation by a new mammalian gene
A;Reference number: A58012; MUID:96178495; PMID:8602223
A;Accession: S68436
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-355,'K',357-519 <DRW>
A;Cross-references: EMBL:U27655
A;Experimental source: tonsil
A;Note: the sequence from Fig. 1b is inconsistent with that from Fig. 1a in having 356-I

```
Query Match          33.2%;  Score 404;  DB 2;  Length 519;
Best Local Similarity 44.2%;  Pred. No. 2.5e-23;
Matches  80;  Conservative  39;  Mismatches  60;  Indels  2;  Gaps  1;

Qy  31 EETSKFAKIR--AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEEAVKWGESFD 88
        ::   ||:|||                     :|||||||||||
Db 337 DEASRKRKSKNLAKDMKNKLGIFRRRNESPGAPPAGKADKMMKSFKPTSEEALKWGESLE 396

Qy  89 KLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAP 148
        ||| ||:|||||||||||
Db 397 KLLVHKYGLAVFQAFLRTEFSEENLEFWLACEDFKKVKSQSKMASKAKKIFAEYIAIQAC 456

Qy 149 KEVNLDFHTKEVITNSITQPTLHSFPDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQR 208
        ||||||| :|              |||||||
Db 457 KEVNLDSYTREHTKDNLQSVTRGCFDLAQKRIFGLMEKDSYPRFLRSDLYLDLINQKKMS 516

Qy 209 P 209
        P
Db 517 P 517
```

RESULT 4
JC7228
G-protein signaling regulator 5 homolog - clawed frog
C;Species: Xenopus sp. (clawed frog)
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C;Accession: JC7228; PC7066
R;Saitoh, O.; Odagiri, M.; Masuho, I.; Nomoto, S.; Kinoshita, N.
Biochem. Biophys. Res. Commun. 270, 34-39, 2000
A;Title: Molecular cloning and characterization of Xenopus RGS5.
A;Reference number: JC7228
A;Accession: JC7228
A;Molecule type: mRNA
A;Residues: 1-181 <SAI>
A;Cross-references: DDBJ:AB038436
A;Accession: PC7066
A;Molecule type: protein
A;Residues: 88-168 <SA2>
A;Experimental source: embryo
C;Comment: This protein has function as a GTPase-activating protein. It interacts with al
genesis and cell differentiation.
C;Superfamily: B-cell activation protein BL34
C;Keywords: differentiation; embryo

```
Query Match          32.9%;  Score 400.5;  DB 2;  Length 181;
Best Local Similarity 47.3%;  Pred. No. 1.3e-23;
Matches  79;  Conservative  35;  Mismatches  50;  Indels  3;  Gaps  1;

Qy  40 RAKEKRNRLSLLVQKPEFHED---TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG 95
        |||     | :||||||
Db  15 RAKEIKTKLGTLLQKPESAIDLIIPYPDKPEKPPKASKPSAEEACQWRDSLEKLLENSYG 74

Qy  96 LEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
        | :| ||||:|||||||||:|||:||  |  ||||||
Db  75 LSVFQSFLKSEFSEENIEFWMACEDYKKAKSPSKMTTKAKKIYEEFIQTEAPREVNIDHF 134

Qy 157 TKEVITNSITQPTLHSFPDAAQSRVVQLMEQDSYTRFLKSDIYLDLME 203
        ||:         |   :| :|||  |||:|| ||
Db 135 TKAVTMKNLVEPSSASFELAQKKIFALMEKDSLPRFVRSEFYQELIK 181
```

RESULT 5
S43436
B cell activation protein BL34 - human
N;Alternate names: B-cell activation protein 1r20; regulator of G-protein signaling 1 (R(
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Jan-2000
C;Accession: S43436; I56165; S34157
R;Newton, J.S.; Deed, R.W.; Mitchell, E.L.D.; Murphy, J.J.; Norton, J.D.
Biochim. Biophys. Acta 1216, 314-316, 1993
A;Title: A B cell specific immediate early human gene is located on chromosome band 1q31
A;Reference number: S43436; MUID:94060109; PMID:8241276
A;Accession: S43436
A;Molecule type: mRNA
A;Residues: 1-196 <NEW>
A;Cross-references: EMBL:X73427; NID:g313214; PIDN:CAA51826.1; PID:g313215
R;Hong, J.X.; Wilson, G.L.; Fox, C.H.; Kehrl, J.H.
J. Immunol. 150, 3895-3904, 1993
A;Title: Isolation and characterization of a novel B cell activation gene.
A;Reference number: I56165; MUID:93232596; PMID:8473738

A;Accession: I56165
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-180,'D',182-196 <HON>
A;Cross-references: GB:S59049; NID:g299704; PID:AAB26289.1; PID:g299705
C;Genetics:
A;Gene: GDB:RGS1; IER1; 1R20; IR20; BL34
A;Cross-references: GDB:439178; OMIM:600323
A;Map position: 1q31-1q31
C;Superfamily: B-cell activation protein BL34
C;Keywords: phosphoprotein

Query Match   28.4%;   Score 346; DB 2; Length 196;
Best Local Similarity   45.9%; Pred. No. 2.1e-19;
Matches   68; Conservative 29; Mismatches 45; Indels 6; Gaps 2;

Qy   55 PEFHEDTRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIE 114
               :||:| |||||: :| :||:|||||||||:||||
Db   46 PHLESGMKSSKSKDV-----LSAAEVMQWSQSLEKLLANQTGQNVFGSFLKSEFSEENIE 100

Qy   115 FWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFD 174
               ||:|||| :||:: ::| ||:|||:|||:|||||||||||:||||
Db   101 FWLACEDYKKTES-DLLPCKAEEIYKAFVHSDAAKQINIDFRTRESTAKKIKAPTPTCFD 159

Qy   175 AAQSRVYQLMEQDSYTRFLKSDIYLDLM 202
               ||||||||||||||
Db   160 EAQKVIYTLMEKDSYPRFLKSHIYLNLL 187

RESULT 6
T21035
hypothetical protein F16H9.1b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C;Accession: T21035; T21272
R;Gardner, A.
submitted to the EMBL Data Library, July 1995
A;Reference number: Z19363
A;Accession: T21035
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-181 <WIL>
A;Cross-references: EMBL:Z50005; PIDN:CAB54219.1; GSPDB:GN00028; CESP:F16H9.1b
R;Gardner, A.
A;Experimental source: clone F16H9
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19398
A;Accession: T21272
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-181 <WI2>
A;Cross-references: EMBL:267882; PIDN:CAB54230.1; GSPDB:GN00028; CESP:F16H9.1b
A;Experimental source: clone F22E10
C;Genetics: CESP:F16H9.1b
A;Map position: X
A;Introns: 28/2; 43/2; 64/1; 103/3
C;Superfamily: B-cell activation protein BL34

Query Match   25.9%;   Score 315.5; DB 2; Length 181;
Best Local Similarity   33.2%; Pred. No. 4e-17;
Matches   67; Conservative 37; Mismatches 67; Indels 31; Gaps 2;

Qy   1 METTLLFPSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
               | :|||:| : :
Db   1 MSCAISCPGKIIVCVTHNSS------PSGKPYVSGSVSVEKKNQEN------------- 40

Qy   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
               :: |: ||||||: |:||:|||||||:||||||||||||||||||||||
Db   41 ---------DGPPTYEIVFGWSQSFENLMKHRAGQKYFAEFLKGEYSDENILFWQACE 89

Qy   121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
               ||| | || || |

Db   90 ELKREKNAEKIEKARIIYEDFISILSPKEVSLDSRVREIVNTNMGRPSASTFDEAQNQI 149

Qy   181 YQLMEQDSYTRFLKSDIYLDLM 202
               | ||||| | ||:||||:|:||
Db   150 YTLMQRDSYPRFLASNIYKTVM 171

RESULT 7
S43576
C05B5.7 protein (clone C05B5) - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 13-Sep-1998
C;Accession: S43576
R;Mortimore, B.
submitted to the EMBL Data Library, April 1994
A;Reference number: S43570
A;Accession: S43576
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-251 <MOR>
A;Cross-references: EMBL:Z32679
C;Genetics:
A;Introns: 165/1; 197/3
C;Superfamily: B-cell activation protein BL34

Query Match   25.7%;   Score 313.5; DB 2; Length 251;
Best Local Similarity   41.5%; Pred. No. 8.4e-17;
Matches   59; Conservative 33; Mismatches 45; Indels 5; Gaps 1;

Qy   66 SGHLAKETRVSPEE------AVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
               | :: || :   :|||| :||::||||:|||||||:||||||||||:::||||
Db   13 SGTVSPQRSVQPEALSYEMVSWQQSFDTLMSFKSGQKCFAEFLKSEYSDENILFWQACE 72

Qy   121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
               ||||||||||||||||||:|||:|||||||||||||:||||||||||||||||||||||
Db   73 ELKREKNSEKMEEKARIIYEDFISILSPKEVSLDSKVREIVNTNMSRPTQNTFEDAQHQI 132

Qy   181 YQLMEQDSYTRFLKSDIYLDLM 202
               || || |||||||:::|:||||
Db   133 YQLMARDSYPRFLTSIFYRETL 154

RESULT 8
G88571
protein C05B5.7 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-Aug-2001
C;Accession: G88571
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele...
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: G88571
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-254 <STO>
A;Cross-references: GB:chr III; PIDN:CAA83595.1; PID:g3873995; GSPDB:GN00021; CESP:C05B5.
A;Note: similar to 1r20 protein (human)
C;Genetics:
A;Gene: C05B5.7
A;Map position: 3
C;Superfamily: B-cell activation protein BL34

Query Match   25.7%;   Score 313.5; DB 2; Length 254;
Best Local Similarity   41.5%; Pred. No. 8.6e-17;
Matches   59; Conservative 33; Mismatches 45; Indels 5; Gaps 1;

Qy   66 SGHLAKETRVSPEE------AVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
               | :: || :   :|||| :||::||||:|||||||:||||||||||:::||||
Db   16 SGTVSPQRSVQPEALSYEMVSWQQSFDTLMSFKSGQKCFAEFLKSEYSDENILFWQACE 75

```
Qy 121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
        |: ||: |||| | ||:||||||||||||| ||||||||||
Db  76 ELKREKNSEKMEEKARIIYEDFISILSPKEVSLDSKVREIVNTNMSRPTQNTFEDAQHQI 135

Qy 181 YQLMEQDSYTRFLKSDIYLDLM 202
        ||||:||||||:|| :| :
Db 136 YQLMARDSYPRFLTSIFYRETL 157
```

RESULT 9
T21034
hypothetical protein F16H9.1a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C;Accession: T21034; T21270
R;Gardner, A.
A;Reference number: Z19363
A;Accession: T21034
submitted to the EMBL Data Library, July 1995
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-169 <WI1>
A;Cross-references: EMBL:Z50005; PIDN:CAA90295.1; GSPDB:GN00028; CESP:F16H9.1a
A;Experimental source: clone F16H9

R;Gardner, A.
A;Reference number: Z19398
A;Accession: T21270
submitted to the EMBL Data Library, November 1995
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-169 <WI2>
A;Cross-references: EMBL:Z67882; PIDN:CAA91803.1; GSPDB:GN00028; CESP:F16H9.1a
A;Experimental source: clone F22E10
C;Genetics:
A;Gene: CESP:F16H9.1a
A;Map position: X
A;Introns: 16/2; 31/2; 52/1; 91/3
C;Superfamily: B-cell activation protein BL34

```
Query Match              25.7%;  Score 313; DB 2; Length 169;
Best Local Similarity    46.7%;  Pred. No. 5.7e-17;
Matches   56;  Conservative   28;  Mismatches   36;  Indels    0;  Gaps    0;

Qy  83 WGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKF 142
        |||| ||||||||||| ||:|||::|||| ||||||:| |:||:|
Db  40 WSQSFENLMKHRAGQKYFAEFLKGEYSDENILFWQACEELKREKNAEKIEEKARIIYEDF  99

Qy 143 IQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLM 202
        ||||||||::||||||||||||||||||||||||||||||||
Db 100 ISILSPKEVSLDSRVREIVNTNMGRPSASTFDEAQNQIYTLMQRDSYPRFLASNIYKTVM 159
```

RESULT 10
T13580
hypothetical protein 52C10.2 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C;Accession: T13580
R;Benos, P.
submitted to the EMBL Data Library, February 1999
A;Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A;Reference number: Z17690
A;Accession: T13580
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-244 <BEN>
A;Cross-references: EMBL:AL035311; PIDN:CAA22951.1
C;Genetics:
A;Cross-references: FlyBase:FBgn0026309
A;Map position: X
A;Introns: 101/2; 127/2; 148/1
A;Note: EG:52C10.2
C;Superfamily: B-cell activation protein BL34

```
Query Match              25.6%;  Score 312; DB 2; Length 244;
Best Local Similarity    48.2%;  Pred. No. 1.1e-16;
Matches   55;  Conservative   23;  Mismatches   36;  Indels    0;  Gaps    0;

Qy  78 EEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKA 137
        ::|:|||| ::: |: |||||||||||| ::|:|||||
Db 131 EEIRSWGKSFDKLMKSTAGRKVFQNFLRSFSEENILFWLACEDLKKENSPELVEEKARL 190

Qy 138 IYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTR 191
        ||||| :||||||||:||||||||||||||||||||||:||||||
Db 191 IYEDYISILSPREVSLDSRVREIVNRNMIEPTHTFDEAQIQIYTLMRDSYPR 244
```

RESULT 11
S71812
RGS10 protein - human
C;Species: Homo sapiens (man)
C;Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 17-Mar-1999
C;Accession: S71812
R;Hunt, T.W.; Fields, T.A.; Casey, P.J.; Peralta, E.G.
Nature 383, 175-177, 1996
A;Title: RGS10 is a selective activator of Galpha(i) GTPase activity.
A;Reference number: S71812; MUID:96371048; PMID:8774883
A;Accession: S71812
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tran
submitted to the EMBL Data Library, November 1995
A;Molecule type: mRNA
A;Residues: 1-173 <HUN>
C;Superfamily: B-cell activation protein BL34

```
Query Match              23.2%;  Score 282.5; DB 2; Length 173;
Best Local Similarity    36.5%;  Pred. No. 1.3e-14;
Matches   62;  Conservative   31;  Mismatches   64;  Indels   13;  Gaps    3;

Qy  56 EFHEDTRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEF 115
        :|:|| || :: || | | |||| ||||::||:|
Db   9 DIHDSDGSSSSSH-----QSLKSTAKWAASLENLLEDPGVKRPREFLKKEFSEENVLF  62

Qy 116 WIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDA 175
        |||||||:|||| ||:|||| :|||||: |||||||||:||||
Db  63 WLACEDFKKMQDKTQMQEKAKEIYMTFLSKASSQVNVEGQSR-LNEKILEEPHPLMFQK 121

Qy 176 AQSRVYQLMEQDSYTRFLKSDIYL------DLMEGRPQRPTNLRRRSRSF 219
        ||||:||:||:|
Db 122 LQDQIFNLMKYDSYSRFLKSDLFLKHKRTEEEEDLPDAQTAAKRASRIY 171
```

RESULT 12
T31002
hypothetical protein F56B6.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 29-Oct-1999
C;Accession: T31002
R;Stellye, L.; Stellye, L.
submitted to the EMBL Data Library, September 1999
A;Description: The sequence of C. elegans cosmid F56B6.
A;Reference number: Z20957
A;Accession: T31002
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-533 <STE>
A;Cross-references: EMBL:U64599; PIDN:AAB04563.1
A;Experimental source: strain Bristol N2
C;Genetics:
A;Map position: X
A;Introns: 39/3; 235/2; 259/2; 322/3; 360/2; 405/1; 461/3; 503/2
A;Note: F56B6.2

```
Query Match              22.9%;  Score 278.5; DB 2; Length 533;
Best Local Similarity    28.8%;  Pred. No. 9.8e-14;
Matches   72;  Conservative   45;  Mismatches   70;  Indels   63;  Gaps    7;

Qy  15 ESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPFHEDT------RSSR  65
```

A;Residues: 1-1387 <SNO>
A;Cross-references: GB:U92280; NID:g2088557; PIDN:AAC53176.1; PID:g2088558
C;Comment: This protein functions as GTPase activating protein. It interacts with ras-lil
F;18-80/Domain: rhophilin-like #status predicted <RHO>
F;712-761/Domain: GH1 #status predicted <GH1>
F;765-800/Domain: GH2 #status predicted <GH2>
F;804-828/Domain: GH3 #status predicted <GH3>
F;1204-1220/Region: conserved #status predicted
F;1266-1295/Region: coiled heptad repeat (S-P-X-S-A)

Query Match   18.5%;   Score 225;   DB 2;   Length 1387;
Best Local Similarity 29.7%;   Pred. No. 3.8e-09;
Matches 68; Conservative 39; Mismatches 92; Indels 30; Gaps 6;

Qy   6 LFFSQINMCESKE-KTFFKLIHGSGKEETSKEAKIRAKEKNRLSLLVQKPEFHED---- 60
              : ||||||| | | : |||
Db   612 LFGPHRNVRKTKEDKKSSKLGRGVALAQTSQRTSAARSFGRSRRFSLTRSLDDLESATVS 671

Qy   61 ------------TRSSRSGHLAKET-----RVSPEEAVKWGESFDKLLSHRDGLEAFTR 102
                         :||:   :|:||
Db   672 DGELTGADLKDCISNNSLSSNASLPSVQSCRRLRERRVASNAVSFERLLQDPVGVRYFSD 731

Qy   103 FLKTEFSEENIEFWIACEDFK--KSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEV 160
              ||:|||||||| || |||||||||||||||||||||||:|||:|:|: ::
Db   732 FLRKEFSEENILFWQACECFSHVPAHDKKELSYRAREIFSKFLCSXATTPVNIDSQA-QL 790

Qy   161 ITNSITQPTLHSFDAAQSRVYQLMEQDSYIRFLKSDIY----LDLMEGR 205
              :| ||:: ||| || || | ||||: || |||
Db   791 ADDILNAPHPDMFKEQQLQIFNLMKFDSYTRFLKSQLYQECVLABVEGR 839

RESULT 15
T19337
hypothetical protein C16C2.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
R;Barlow, K.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19109
A;Accession: T19337
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-473 <WIL>
A;Cross-references: EMBL:Z81036; PIDN:CAB02742.1; GSPDB:GN00019; CESP:C16C2.2
A;Experimental source: clone C16C2
C;Genetics:
A;Gene: CESP:C16C2.2
A;Map position: 1
A;Introns: 7/3; 47/3; 75/3; 107/1; 133/2; 232/2; 270/2; 440/3; 469/2

Query Match   18.1%;   Score 220;   DB 2;   Length 473;
Best Local Similarity 33.7%;   Pred. No. 2.5e-09;
Matches 59; Conservative 29; Mismatches 67; Indels 20; Gaps 5;

Qy   57 FHEDTRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFW 116
              | ::|| |||||||||||||||||||||||||||| :
Db   263 FWNQPTDTSSAEIPTBKRVK----RWGLSVQELVKDPIGRQVLETFLESEFSSENIRFW 317

Qy   117 IACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAA 176
              ||| | ||| ||||||| :| : :::|
Db   318 IAIQDLKYAPN-EQIVQKAERIREEFLAQGAPAQVNVDNRTLDQTLECISK---AKDAS 372

Qy   177 QSR--------VYQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCN 222
              | ||||||||:|||:
Db   373 QMRFAFYHSEEHVFTLMAKDSYPRFVRSQIYKAVLTAAQQHGTK-RLGWRNVFN 426

Search completed: August 20, 2004, 16:47:48
Job time : 27 secs

Db   299 EINNEGFPLLNKDQGRKQNFPQRKVR-----------RQTFYEDPAFTGVSSAGGSVI 345
              : || |||:   : |||
Qy   66 SGHLAKETRV-----------------------SPEEAVKWGESFDKLLSHRD 95
              :|:||||:   :||
Db   346 SNNTGGQVRIALKVIGPISRTLSYLRSKMDLALSTSSLYPSRDDVRQWEISFESLLNNKF 405

Qy   96 GLEAFTRFLKTEFSEENIEFWIACEDFKKSK-GPQQIHLKAKAIYEKFIQTDAPKEVNLD 154
              ||||| |:||||||||| | | : | ||||||||||||||| :||: ::||
Db   406 GCALFRQFLKKEFSDENMDFWLECEEFKFKMKDGKKKSTTQKAIEIYSEFVAEHSPKEVNLD 465

Qy   155 FHTKEVITNSI---TQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEG------RP 206
              ||||: |: : ||::| || |||||||||||: ||:||| |
Db   466 SDTRAATKAAVEAGCKP--DTFALAQSRVEQLMSKDSYRRFLRDRLFLDLLESYEITDKE 523

Qy   207 QRPTNLRRRS 216
              :| |: ::
Db   524 DKPSSSKDKN 533

RESULT 13
JC5503
G-protein signaling regulator 14 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Sep-1997 #sequence_revision 05-Nov-1999
C;Accession: JC5503
R;Snow, B.E.; Antonio, L.; Suggs, S.; Gutstein, H.B.; Siderovski, D.P.
Biochem. Biophys. Res. Commun. 233, 770-777, 1997
A;Title: Molecular cloning and expression analysis of rat Rgs12 and Rgs14.
A;Reference number: JC5502; MUID:97312490; PMID:9168931
A;Accession: JC5503
A;Molecule type: mRNA
A;Residues: 1-544 <SNO>
A;Cross-references: GB:U92279; NID:g2088555; PIDN:AAC53175.1; PID:g2088556
C;Comment: This protein functions as GTPase activating protein.
F;64-113/Domain: GH1 #status predicted <GH1>
F;117-152/Domain: GH2 #status predicted <GH2>
F;155-180/Domain: GH3 #status predicted <GH3>
F;514-534/Region: conserved #status predicted

Query Match   20.0%;   Score 243.5;   DB 2;   Length 544;
Best Local Similarity 34.6%;   Pred. No. 4.7e-11;
Matches 63; Conservative 30; Mismatches 62; Indels 27; Gaps 6;

Qy   59 EDTRSSRSGHLAKETRVSP-----EEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENI 113
              |||:|||||||||||| |:||||||||||||||||||||||||||||||||||
Db   35 EGRGSLSIHSLPSGPSSPFSTDEQPVASWAQSFERLLQDPRGLAYFTEFLKKEFSAENV 94

Qy   114 EFWIACEDFKK--SKGPQQIHLKAKAIYEKFIQTDAPKEVNLD---FHTKEVITNSITQP 168
              |||| | | |||||||||||||||||||:   : :|| | ::||:
Db   95 TFWQACERFQQIPASDTKQLAQEAHNIYHEFLSSQALSPVNIDRQAWLSBEV----LAQP 150

Qy   169 TLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLM----EGRPQR-------PTNLRRR 215
              ||:||||||||||||||||:||||||:|||||:   :| |:|:|||
Db   151 RPDMFRAQQLQIFNLMKFDSYARFVKSPLYQECLLAEEAGRPLREPGSSHLGSPDTARKK 210

Qy   216 SR 217
              ||
Db   211 PK 212

RESULT 14
JC5502
G-protein signaling regulator 12 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 05-Nov-1999
C;Accession: JC5502
R;Snow, B.E.; Antonio, L.; Suggs, S.; Gutstein, H.B.; Siderovski, D.P.
Biochem. Biophys. Res. Commun. 233, 770-777, 1997
A;Title: Molecular cloning and expression analysis of rat Rgs12 and Rgs14.
A;Reference number: JC5502; MUID:97312490; PMID:9168931
A;Accession: JC5502
A;Molecule type: mRNA

```
                    GenCore version 5.1.6
         Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    August 20, 2004, 12:32:00 ; Search time 14.5 Seconds
                                (without alignments)
                     843.896 Million cell updates/sec

Title:           US-09-894-749-2
Perfect score:   1218
Sequence:        1 METLLLFFSQINMCESKEKT.........SRSFTCNEFQDVQSDVAIWL 235

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters:        141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_42:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                           SUMMARIES
```

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1218	100.0	235	1	RGSI_HUMAN	Q9ns28 homo sapien
2	1032	84.7	235	1	RGSI_MOUSE	Q99pg4 mus musculu
3	422.5	34.7	211	1	RGS2_HUMAN	P41220 homo sapien
4	421.5	34.6	181	1	RGS5_HUMAN	O15539 homo sapien
5	418.5	34.4	211	1	RGS2_RAT	Q9jhx0 rattus norv
6	416.5	34.2	211	1	RGS2_MOUSE	O08849 mus musculu
7	410.5	33.7	181	1	RGS5_RAT	P49800 rattus norv
8	409	33.6	568	1	RGS3_MOUSE	Q9dc04 mus musculu
9	408.5	33.5	205	1	RGS4_HUMAN	P49798 homo sapien
10	405.5	33.3	205	1	RGS4_MOUSE	O08899 mus musculu
11	404.5	33.2	181	1	RGS5_MOUSE	O08850 mus musculu
12	404.5	33.2	205	1	RGS4_RAT	P49799 rattus norv
13	404	33.2	519	1	RGS3_HUMAN	P49796 homo sapien
14	402.5	33.0	180	1	RGS8_HUMAN	P57771 homo sapien
15	402.5	33.0	180	1	RGS8_RAT	P49804 rattus norv
16	380.5	31.2	196	1	RGS1_MOUSE	Q9j125 mus musculu
17	378	31.0	202	1	RGSG_HUMAN	O15492 homo sapien
18	378	31.0	202	1	RGSG_BOVIN	O46471 bos taurus
19	369	30.3	201	1	RGSG_MOUSE	P97428 mus musculu
20	366.5	30.1	199	1	RGSG_RAT	P56700 rattus norv
21	353	29.0	196	1	RGS1_HUMAN	Q08116 homo sapien
22	327.5	26.9	218	1	RGSK_CHICK	Q9wal gallus gall
23	325	26.7	388	1	RGSK_HUMAN	O76081 homo sapien
24	324	26.6	239	1	RGSK_MOUSE	Q9qzb1 mus musculu
25	318.5	26.1	374	1	RGSK_BOVIN	P79348 bos taurus
26	315	25.9	210	1	RGSH_MOUSE	Q9qzb0 mus musculu
27	313.5	25.7	201	1	RGSI_CAEEL	P34295 caenorhabdi
28	313	25.7	169	1	RGS2_CAEEL	P49808 caenorhabdi
29	311.5	25.6	216	1	RGSJ_RAT	O70521 rattus norv
30	309.5	25.4	210	1	RGSH_CHICK	Q9wa0 gallus gall
31	309.5	25.4	216	1	RGSJ_MOUSE	Q9cx84 mus musculu
32	308.5	25.3	210	1	RGSH_HUMAN	Q9ugc6 mus musculu
33	306	25.1	217	1	RGSJ_HUMAN	P49795 homo sapien

```
34 297.5 24.4  RGSA_MOUSE  181 1  Q9cge5 mus musculu
35 285   23.4  RGSD_HUMAN  159 1  O14921 homo sapien
36 282.5 23.2  RGSA_HUMAN  173 1  O43665 homo sapien
37 270.5 22.2  RGS6_MOUSE  409 1  Q9zzh2 mus musculu
38 261.5 21.5  RGS7_HUMAN  495 1  P49802 homo sapien
39 256   21.0  RGS7_RAT    477 1  P49803 rattus norv
40 254   20.9  RGS7_BOVIN  469 1  O46470 bos taurus
41 252.5 20.7  RGS9_HUMAN  674 1  O75916 homo sapien
42 251   20.6  RGS7_MOUSE  469 1  O54829 mus musculu
43 248.5 20.4  RGS9_MOUSE  675 1  O54828 mus musculu
44 247.5 20.3  RGS9_RAT    677 1  P49805 rattus norv
45 245.5 20.2  RGS9_BOVIN  484 1  O46469 bos taurus
```

```
                              ALIGNMENTS

RESULT 1
RGSI_HUMAN
ID   RGSI_HUMAN      STANDARD;      PRT;     235 AA.
AC   Q9ns28;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Regulator of G-protein signaling 18 (RGS18).
GN   RGS18 OR RGS13.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhang W., Wan T., Yuan Z., He L., Cao X.;
RT   "A novel regulator of G-protein signaling.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Platelet;
RA   Gagnon A.W., Murray D.L., Leadley R.J. Jr.;
RT   "Cloning and characterization of a novel regulator of G-protein
RT   signaling in human platelets";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   PubMed=11042171;
RA   Park I.K., Klug C.A., Li K., Jerabek L., Li L., Nanamori M.,
RA   Neubig R.R., Hood L., Weissman I.L., Clarke M.F.;
RT   "Molecular cloning and characterization of a novel regulator of
RT   G-protein signaling from mouse hematopoietic stem cells.";
RL   J. Biol. Chem. 276:915-923 (2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
```

```
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. BINDS TO G(I)-ALPHA AND G(Q)-ALPHA
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in peripheral leukocytes, bone
CC       marrow, spleen and fetal liver.
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF076642; AAF80227.1; -.
DR   EMBL; AF268036; AAK58589.1; -.
DR   EMBL; BC020632; AAH20632.1; -.
DR   HSSP; P49799; 1AGR.
DR   Genew; HGNC:14261; RGS18.
DR   MIM; 607192; -.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor. RGS.
FT   DOMAIN        86    202       RGS.
FT   CONFLICT     226    227       DV -> ML (IN REF. 3).
SQ   SEQUENCE   235 AA;   27582 MW;  973ABDE8EC7DE3D5 CRC64;

Query Match             100.0%;  Score 1218;  DB 1;  Length 235;
Best Local Similarity   100.0%;  Pred. No. 3.4e-85;
Matches  235;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235

RESULT 2
RGSI_MOUSE
ID   RGSI_MOUSE     STANDARD;      PRT;   235 AA.
AC   Q99PG4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Regulator of G-protein signaling 18 (RGS18).
GN   RGS18.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BA;
RX   PubMed=11042171;
```

```
RA   Park I.K., Klug C.A., Li K., Jerabek L., Li L., Nanamori M.,
RA   Neubig R.R., Hood L., Weissman I.L., Clarke M.F.;
RT   "Molecular cloning and characterization of a novel regulator of
RT   G-protein signaling from mouse hematopoietic stem cells.";
RL   J. Biol. Chem. 276:915-923(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BDF1;
RX   MEDLINE=21240406; PubMed=11342430;
RA   Nagata Y., Oda M., Nakata H., Shozaki Y., Kozasa T., Todokoro K.;
RT   "A novel regulator of G-protein signaling bearing GAP activity for
RT   Galphai and Galphaq in megakaryocytes.";
RL   Blood 97:3051-3060(2001).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. BINDS TO G(I)-ALPHA AND G(Q)-ALPHA.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: Expressed in bone marrow, spleen, fetal liver
CC       and lung. At very low levels expressed in heart.
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF302685; AAG53657.1; -.
DR   EMBL; AB042807; BAB60723.1; -.
DR   HSSP; P49799; 1AGR.
DR   MGD; MGI:1927498; Rgs18.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005096; F:GTPase activator activity; IDA.
DR   GO; GO:0016299; F:regulator of G-protein signaling activity; IDA.
DR   GO; GO:0008277; P:regulation of G-protein coupled receptor pr... .; IDA.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor. RGS.
FT   DOMAIN        86    202       RGS.
SQ   SEQUENCE   235 AA;   27610 MW;  6433435AC1C8CFD8 CRC64;

Query Match              84.7%;  Score 1032;  DB 1;  Length 235;
Best Local Similarity    83.8%;  Pred. No. 3.8e-71;
Matches  197;  Conservative   19;  Mismatches   19;  Indels    0;  Gaps    0;

Qy    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60
        |::||:||||||||||||||||||:||||||||||:||||||||||||||||:||::|||
Db    1 MDMSLVFFSQINMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLQRPDFHGE 60

Qy   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120
        |::||:|:|||||||||||||||:|||||||||||||:|||||||||||||||||||:||
Db   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120

Qy  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180
        ||||.|.|||||:|||||||||||.||||||||:|||||||.||:|||||||||||:|||
Db  121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180

Qy  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
        |||||:|||.|||||.:||.||||||||||||||||||||.:||:|||:||||||
Db  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235

RESULT 3
RGS2_HUMAN
ID   RGS2_HUMAN     STANDARD;      PRT;   211 AA.
AC   P41220;
DT   01-FEB-1995 (Rel. 31, Created)
```

```
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Regulator of G-protein signaling 2 (RGS2) (G0/G1 switch regulatory
DE   protein 8).
GN   RGS2 OR G0S8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=94235158; PubMed=8179820;
RA   Siderovski D.P., Heximer S.P., Forsdyke D.R.;
RT   "A human gene encoding a putative basic helix-loop-helix
RT   phosphoprotein whose mRNA increases rapidly in cycloheximide-treated
RT   blood mononuclear cells.";
RL   DNA Cell Biol. 13:125-147(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
RT   "cDNA clones of human proteins involved in signal transduction
RT   sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Bagguley C.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Urinary bladder;
RX   MEDLINE=22338257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   CHARACTERIZATION.
RX   MEDLINE=95371353; PubMed=7643615;
RA   Wu H.-K., Heng H.H., Shi X.-M., Forsdyke D.R., Tsui L.-C., Mak T.W.,
RA   Minden M.D., Siderovski D.P.;
RT   "Differential expression of a basic helix-loop-helix phosphoprotein
RT   gene, G0S8, in acute leukemia and localization to human chromosome
RT   1q31.";
RL   Leukemia 9:1291-1298(1995).
RN   [6]
RP   PHOSPHORYLATION.
RX   MEDLINE=21101914; PubMed=11063746;
RA   Cunningham M.L., Waldo G.L., Hollinger S., Hepler J.R., Harden T.K.;
RT   "Protein kinase C phosphorylates RGS2 and modulates its capacity for
RT   negative regulation of Galpha i1 signaling.";
RL   J. Biol. Chem. 276:5438-5444(2001).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. MAY PLAY A ROLE IN LEUKEMOGENESIS.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ACUTE MYELOGENOUS LEUKEMIA (AML)
CC       AND IN ACUTE LYMPHOBLASTIC LEUKEMIA (ALL).
CC   -!- PTM: Phosphorylated by protein kinase C.
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L13391; AAA20680.1; -.
DR   EMBL; L13463; AAC37587.1; -.
DR   EMBL; AF493926; AAM12640.1; -.
DR   EMBL; AL035407; CAB62512.1; -.
DR   EMBL; BC007049; AAH07049.1; -.
DR   PIR; I53020; I53020.
DR   HSSP; P49799; 1AGR.
DR   Genew; HGNC:9998; RGS2.
DR   MIM; 600861; -.
DR   GO; GO:0005516; F:calmodulin binding; TAS.
DR   GO; GO:0005096; F:GTPase activator activity; TAS.
DR   GO; GO:0008277; P:regulation of G-protein coupled receptor pr...; TAS.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Cell cycle; Phosphorylation.
FT   DOMAIN       83    199       RGS.
SQ   SEQUENCE   211 AA;  24382 MW;  EFFF4AE47EF9AD8F CRC64;

Query Match             34.7%;  Score 422.5;  DB 1;  Length 211;
Best Local Similarity   42.8%;  Pred. No. 3.4e-25;
          Matches  89;  Conservative  37;  Mismatches  73;  Indels  9;  Gaps  3;

Qy      6 LFFSQINMCESKEKTFKLIHGSGKEETSKEAKIR---AKEKRNRLSLLVQKPEFHEDTR 62
            : |: |  | | |:|    ||   || |:   ||:   |  | ||:|| | | :||
Db      5 MFLAVQHDCRPMDKS----AGSGHKSEEKREKMKRTLLKDWKTRLSYFIQNSSTPGKPK 59

Qy     63 SSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRPLKTEFSEENIEFWIACEDF 122
            |:   |||:|:|||  ||: |: |:|||||:||||:  |||:||:||:|||||||:|||
Db     60 TGKKSKQQAFIKPSPEEAQLWSEAFDLLASKYGLAAFRAFLKSEFCEENIEFWLACEDF 119

Qy    123 KKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVIQ 182
           |||| :  :  || |:| ||  |: :|: ||: | :|| |:||::|||:||:|| ||::|
Db    120 KKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQBATSGCFTTAQXRVYS 179

Qy    183 LMEQDSYTRFLKSDIYLDLMEGRPQRPT 210
           |||:||||||||||:|| : |:|:|
Db    180 LMENNSYPRFLESEFYQDLCK-KPQITT 206

RESULT 4
RGS5_HUMAN
ID   RGS5_HUMAN     STANDARD;      PRT;    181 AA.
AC   O15539;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Regulator of G-protein signaling 5 (RGS5).
GN   RGS5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Chatterjee T.K., Fisher R.A.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Neuroblastoma;
RX   MEDLINE=98419174; PubMed=9747037;
RA   Seki N., Sugano S., Suzuki Y.-I., Nakagawara A., Ohira M.,
RA   Muramatsu M.-A., Saito T., Hori T.;
RT   "Isolation, tissue expression, and chromosomal assignment of human
RT   RGS5, a novel G-protein signaling regulator gene.";
RL   J. Hum. Genet. 43:202-205(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=99403338; PubMed=10471929;
RA   Cismowski M.J., Takesono A., Ma C., Lizano J.S., Xie X.,
RA   Fuernkranz H., Lanier S.M., Duzic E.;
RT   "Genetic screens in yeast to identify mammalian nonreceptor modulators
RT   of G-protein signaling.";
RL   Nat. Biotechnol. 17:878-883(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
RT   "cDNA clones of human proteins involved in signal transduction
RT   sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. BINDS TO G(I)-ALPHA AND G(O)-ALPHA,
CC       BUT NOT TO G(S)-ALPHA (BY SIMILARITY).
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF030108; AAB84001.1; -.
DR   EMBL; AB008109; BAA22889.1; -.
DR   EMBL; AF159570; AAD40957.1; -.
DR   EMBL; AF493929; AAM12643.1; -.
DR   EMBL; BC030059; AAH30059.1; -.
DR   HSSP; P49799; 1AGR.
DR   Genew; HGNC:10001; RGS5.
DR   MIM; 603276; -.
DR   GO; GO:0005096; F:GTPase activator activity; TAS.
DR   GO; GO:0008277; P:regulation of G-protein coupled receptor pr...; TAS.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.

DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor.
FT   DOMAIN       64    180       RGS.
SQ   SEQUENCE   181 AA;   20946 MW;  2E08CB0179DB7687 CRC64;

Query Match             34.6%;  Score 421.5;  DB 1;  Length 181;
Best Local Similarity   49.1%;  Pred. No. 3.3e-25;
Matches   82;  Conservative   37;  Mismatches   45;  Indels    3;  Gaps    1;

Qy    40 RAKEKRNRLSLLVQKPEFHED---TRSSRSGHLAKETRVSPERAVKWGESFDKLLSHRDG 96
          |||| ||:||||:||| |::   :|||: :||:::   ||| : :|||||:|||||:|:
Db    15 RAKEIKIKLGLLQKPDSVGDLVIPYNEKPKPAKTQKTSLDEALQWRDSLDKLLQNNYG 74

Qy    97 LEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
          | :|: |||:||||||:|||||||:||  | |: |:||:|| ||||||| |||||||:| 
Db    75 LASFKSFLKSEFSEENLEFWIACEDYKKIKSPAKMAEKAKQIYBEFIQTEAPKEVNIDHF 134

Qy   157 TKEVITNSITQPTLHSFDAAQSRVVQLMEQDSYTRFLKSDIYLDLME 203
          ||::| :|:|::|:||| ||: :||| :||||| |:| :: :|  :
Db   135 TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK 181


RESULT 5
RGS2_RAT
ID   RGS2_RAT      STANDARD;      PRT;   211 AA.
AC   Q9UHX0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Regulator of G-protein signaling 2 (RGS2).
GN   RGS2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Thoracic aorta;
RX   MEDLINE=20159060; PubMed=10692485;
RA   Grant S.L., Lassegue B., Griendling K.K., Ushio-Fukai M., Lyons P.R.,
RA   Alexander R.W.;
RT   "Specific regulation of RGS2 messenger RNA by angiotensin II in
RT   cultured vascular smooth muscle cells.";
RL   Mol. Pharmacol. 57:460-467(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Brain;
RA   Taymans J.-M., Wintmolders C., Te Riele P., Jurzak M.,
RA   Groenewegen H.J., Leysen J.E., Langlois X.;
RT   "A detailed distribution study of RGS2 messenger RNA and protein in
RT   the rat brain.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Ko J.K., Kim I.S., Park D.H.;
RT   "PTH induced RGS cDNA sequence in rat osteoblast-like UMR106 cell.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 1-100 FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=21302414; PubMed=11409749;
RA   Robinet E.A., Wurch T., Pauwels P.J.;
RT   "Haloperidol and clozapine differently affect RGS2 mRNA expression in
RT   rat brain.";
RL   NeuroReport 12:1731-1735(2001).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. MAY PLAY A ROLE IN LEUKEMOGNESIS
CC       (BY SIMILARITY).
CC   -!- PTM: Phosphorylated by protein kinase C (By similarity).
CC   -!- SIMILARITY: Contains 1 RGS domain.
```

```
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF279918; AAF85981.1; -.
DR   EMBL; AF321837; AAK09375.1; -.
DR   EMBL; AY043246; AAK85309.1; -.
DR   EMBL; AJ318489; CAC44900.1; -.
DR   HSSP; P49799; 1AGR.
DR   InterPro; IPR000342; Reg1_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Reg1_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Cell cycle; Phosphorylation.
FT   DOMAIN       83    199       RGS.
SQ   SEQUENCE   211 AA;  24323 MW;  84B5525CB41EA964 CRC64;

Query Match             34.4%;  Score 418.5;  DB 1;  Length 211;
Best Local Similarity   42.3%;  Pred. No. 6.7e-25;
Matches    88;  Conservative    39;  Mismatches    72;  Indels     9;  Gaps     3;

Qy      6 LFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTR 62
            :|: ||  ::   ::|:::  :|:||:  |||  :|::||    |||:  |||||
Db      5 MFLAVQHDCVPMDKS-----AGNGPKVEEKREKMKRTLLKDWKTRLSYFLQNSSTPGKPK 59

Qy     63 SSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDF 122
            |:| ||| :|: ||: ||:||   :|||:|   :|| || || :|  :|||||||||||
Db     60 TGKKSKQQTFIKPSPEEALLWAEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACEDF 119

Qy    123 KKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQ 182
            |||| || ||  ||| :  :||||| |:| |||: |: |  :|||||||| |||  :|
Db    120 KKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKLIAQNIQEATSGCFTTAQKRVYS 179

Qy    183 LMEQDSYTRFLKSDIYLDLMEGRPQRPT 210
            ||||||||| ||: |: ::||||||
Db    180 LMENNSYPRFLESEFYQDLCK-KPQITT 206

RESULT 6
RGS2_MOUSE
ID   RGS2_MOUSE     STANDARD;      PRT;    211 AA.
AC   O08849; Q91WX1; Q9JL24;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Regulator of G-protein signaling 2 (RGS2).
GN   RGS2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97236828; PubMed=9079700;
RA   Chen C., Zheng B., Han J., Lin S.C.;
RT   "Characterization of a novel mammalian RGS protein that binds to
RT   Galpha proteins and inhibits pheromone signaling in yeast.";
RL   J. Biol. Chem. 272:8679-8665(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=20243574; PubMed=10779778;
RA   Reif K., Cyster J.G.;
RT   "RGS molecule expression in murine B lymphocytes and ability to
RT   down-regulate chemotaxis to lymphoid chemokines.";
RL   J. Immunol. 164:4720-4729(2000).
```

```
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RA   Shen Q.-X., Wang J., Huang Z.-P.;
RT   "Identification of novel endometrial and embryonic factors involved in
RT   mouse embryo implantation.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
    ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
    THEIR INACTIVE GDP-BOUND FORM. MAY PLAY A ROLE IN LEUKEMOGENESIS.
-!- TISSUE SPECIFICITY: Expressed in a wide variety of tissues.
-!- PTM: Phosphorylated by protein kinase C (By similarity).
-!- SIMILARITY: Contains 1 RGS domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U67187; AAB50617.1; -.
DR   EMBL; AF215668; AAF34625.1; -.
DR   EMBL; AF432916; AAL28114.1; -.
DR   HSSP; P49799; 1AGR.
DR   MGD; MGI:1098271; Rgs2.
DR   InterPro; IPR000342; Reg1_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Reg1_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Cell cycle; Phosphorylation.
FT   DOMAIN       83    199       RGS.
FT   CONFLICT     39     39       KD -> NH (IN REF. 1).
FT   CONFLICT     78     79       QL -> HV (IN REF. 1).
FT   CONFLICT     78     79       QL -> DV (IN REF. 3).
SQ   SEQUENCE   211 AA;  24294 MW;  5D6E255C2BC7E7FA CRC64;

Query Match             34.2%;  Score 416.5;  DB 1;  Length 211;
Best Local Similarity   42.3%;  Pred. No. 9.5e-25;
Matches    88;  Conservative    38;  Mismatches    73;  Indels     9;  Gaps     3;

Qy      6 LFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTR 62
            :|: ||  ::   ::|:::  :|:||:  |||  :|::||    |||:  |||||
Db      5 MFLAVQHDCVPMDKS-----AGNGPKVEEKREKMKRTLLKDWKTRLSYFLQNSSAPGKPK 59

Qy     63 SSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDF 122
            |:| ||| :|: ||: ||:||   :|||:|   :|| || || :|  :|||||||||||
Db     60 TGKKSKQQTFIKPSPEEAQLWAEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACEDF 119

Qy    123 KKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQ 182
            |||| || ||  ||| :  :||||| |:| |||: |: |  :|||||||| |||  :|
Db    120 KKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKLIAQNIQEATSGCFTTAQKRVYS 179

Qy    183 LMEQDSYTRFLKSDIYLDLMEGRPQRPT 210
            ||||||||| ||: |: ::||||||
Db    180 LMENNSYPRFLESEFYQDLCK-KPQITT 206

RESULT 7
RGS5_RAT
ID   RGS5_RAT       STANDARD;      PRT;    181 AA.
AC   P49800; Q93KD7;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Regulator of G-protein signaling 5 (RGS5).
GN   RGS5.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
```

```
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kirsch T., Lippoldt A., Wellner M., Haller H.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 95-161 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=96140645; PubMed=8548815;
RA   Koelle M.R., Horvitz H.R.;
RT   "EGL-10 regulates G protein signaling in the C. elegans nervous
RT   system and shares a conserved domain with many mammalian proteins.";
RL   Cell 84:115-125(1996).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. BINDS TO G(I)-ALPHA AND G(O)-ALPHA,
CC       BUT NOT TO G(S)-ALPHA (BY SIMILARITY).
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF241159; AAF73424.1; -.
DR   EMBL; U32435; AAC52372.1; -.
DR   HSSP; P49799; 1AGR.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; RGS.
FT   DOMAIN        64    180       RGS.
SQ   SEQUENCE   181 AA;  21052 MW;  9C179CE008B82C03 CRC64;

Query Match             33.7%;  Score 410.5;  DB 1;  Length 181;
Best Local Similarity   48.2%;  Pred. No. 2.2e-24;
Matches    80;  Conservative    35;  Mismatches   48;  Indels    3;  Gaps    1;

Qy      40 RAKEKRNRLSLLVQKPEFHED--TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG 96
              ||||| :|| :|::  :|:  :|||:: ||  : || :|| : :|||  ||  ||
Db      15 RAKEIKILGILLQKPDSAVDLVIPYNEKPEKPAKAHKPSLEEVLQWRQSLDKLLQSNYG 74

Qy      97 LEAFTRFLKTEFSEENIEFWIACEDFKKSKGPGQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
             |||  |||| ||:|||| | ||| |::  ||  ||| | ||  |:|| |: ||||||| :|
Db      75 FASFKSFLKSEFSEENLEFWVACENYKKIKSPIKVAEKAKQIYEBFIQTEAPKEVNIDHF 134

Qy     157 TKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLM 202
              :::  |::|:|| |||| ||  :|| ||:|::| :|:
Db     135 TKDITMKNLVEPSPHSFDLAQKRIYALMEKDSLPRFVRSEFYKELI 180

RESULT 8
ID   RGS3_MOUSE     STANDARD;      PRT;   568 AA.
AC   Q9DC04; Q9UL22; Q9UL23;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Regulator of G-protein signaling 3 (RGS3).
GN   RGS3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM LONG)
RC   STRAIN=C57BL/6J; TISSUE=Lung;
```

```
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King D., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM LONG)
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Matsina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   PARTIAL SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=20243574; PubMed=10779778;
RA   Reif K., Cyster J.G.;
RT   "RGS molecule expression in murine B lymphocytes and ability to
RT   down-regulate chemotaxis to lymphoid chemokines.";
RL   J. Immunol. 164:4720-4729(2000).
CC   -!- FUNCTION: Inhibits signal transduction by increasing the GTPase
CC       activity of G protein alpha subunits thereby driving them into
CC       their inactive GDP-bound form.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Long;
CC         IsoId=Q9DC04-1; Sequence=Displayed;
CC       Name=Short;
CC         IsoId=Q9DC04-2; Sequence=VSP_005663, VSP_005664;
CC   -!- PTM: Phosphorylated by cyclic GMP-dependent protein kinase (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC    ------------------------------
DR    EMBL; AK004648; BAB23439.1; -.
DR    EMBL; BC033449; AAH33449.1; -.
DR    EMBL; AF215670; AAF34627.1; -.
DR    EMBL; AF215669; AAF34626.1; -.
DR    HSSP; P49799; 1AGR.
DR    MGD; MGI:1354734; Rgs3.
DR    GO; GO:0005634; C:nucleus; IDA.
DR    InterPro; IPR000342; Regl_Gprotein.
DR    Pfam; PF00615; RGS; 1.
DR    PRINTS; PR01301; RGSPROTEIN.
DR    ProDom; PD001580; Regl_Gprotein; 1.
DR    SMART; SM00315; RGS; 1.
DR    PROSITE; PS50132; RGS; 1.
KW    Signal transduction inhibitor; Alternative splicing; Phosphorylation.
FT    DOMAIN      443    559      RGS.
FT    VARSPLIC      1    375      Missing (in isoform Short).
FT                               /FTId=VSP_005663.
FT    VARSPLIC    376    397      KMSGTDLTECDEASRKRKRSKNI -> MLRGMYLTRNGNLQR
FT                               RHTMKE (in isoform Short).
FT                               /FTId=VSP_005664.
SQ    SEQUENCE    568 AA;  61558 MW;  91939E22F3DA9F5C CRC64;

Query Match             33.6%;  Score 409;  DB 1;  Length 568;
Best Local Similarity   45.3%;  Pred. No. 1.2e-23;
Matches   82;  Conservative  37;  Mismatches  60;  Indels   2;  Gaps    1;

Qy     31  EETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEEAVKWGESFD  88
           :| | ||:|:| ||:.|:|::  :: ::   ::|:   ||:.:.::|||:||||:| |||
Db    386  DEASRKRKSKNIAKDMKNKLAIPRRRNESPGAQPASKTDKTTKSFKPTSERALKWSESLE  445

Qy     89  KLLSHRDGLEAFTRFLKTEFSEENIEPWIACEDFKKKSKGPQQIHLKAKAIYEKFIQTDAP 148
           |||: : || ||:|:|:||| || || ||||||:|| : ||:|.||:| |:|||:| ::||
Db    446  KLLLHKYGLEVFQAFLRTEFSEENLEFWLACEDFKKVKSQSKWAAKAKKIFAEFIAIQAC  505

Qy    149  KEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQR  208
           ||||||| | :| :|:|  || || |:|::|:|| :|:||:|||| |::|:|||:|||:
Db    506  KEVNLDSYTREHTKENLQSITRGCFDLAQKRIFGLMEKDSYPRFLRSDLYLDLINQKKMS  565

Qy    209  P  209
           |
Db    566  P  566
```

RESULT 9
```
RGS4_HUMAN
ID    RGS4_HUMAN     STANDARD;     PRT;    205 AA.
AC    P49798;
DT    01-OCT-1996 (Rel. 34, Created)
DT    01-OCT-1996 (Rel. 34, Last sequence update)
DT    15-MAR-2004 (Rel. 43, Last annotation update)
DE    Regulator of G-protein signaling 4 (RGS4) (RGP4).
GN    RGS4.
OS    Homo sapiens (Human).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX    NCBI_TaxID=9606;
RN    [1]
RP    SEQUENCE FROM N.A.
RX    MEDLINE=96178495; PubMed=8602223;
RA    Druey K.M., Blumer K.J., Kang V.H., Kehrl J.H.;
RT    "Inhibition of G-protein-mediated MAP kinase activation by a new
RT    mammalian gene family.";
RL    Nature 379:742-746(1996).
RN    [2]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Brain, and Uterus;
RA    Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
RT    "cDNA clones of human proteins involved in signal transduction
RT    sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL    Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN    [3]
RP    SEQUENCE FROM N.A.
```

```
RC    TISSUE=Brain;
RX    MEDLINE=22388257; PubMed=12477932;
RA    Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA    Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA    Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA    Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA    Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA    Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA    Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA    Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA    Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA    Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA    Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA    Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA    Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA    Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA    Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA    Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA    Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT    "Generation and initial analysis of more than 15,000 full-length
RT    human and mouse cDNA sequences.";
RL    Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN    [4]
RP    PALMITOYLATION.
RX    MEDLINE=20076507; PubMed=10608901;
RA    Tu Y., Popov S., Slaughter C., Ross E.M.;
RT    "Palmitoylation of a conserved cysteine in the regulator of G protein
RT    signaling (RGS) domain modulates the GTPase-activating activity of
RT    RGS4 and RGS10.";
RL    J. Biol. Chem. 274:38260-38267(1999).
RN    [5]
RP    INHIBITION.
RX    MEDLINE=98421527; PubMed=9748280;
RA    Wang J., Ducret A., Tu Y., Kozasa T., Aebersold R., Ross E.M.;
RT    "RGSZ, a Gz-selective RGS protein in brain. Structure, membrane
RT    association, regulation by Galphaz phosphorylation, and relationship
RT    to a Gz GTPase-activating protein subfamily.";
RL    J. Biol. Chem. 273:26014-26025(1998).
CC    -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC        ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC        THEIR INACTIVE GDP-BOUND FORM. ACTIVITY ON G(Z)-ALPHA IS INHIBITED
CC        BY PHOSPHORYLATION OF THE G-PROTEIN. ACTIVITY ON G(Z)-ALPHA AND
CC        G(I)-ALPHA-1 IS INHIBITED BY PALMITOYLATION OF THE G-PROTEIN.
CC    -!- PTM: Either Cys-2 or Cys-12 or both are palmitoylated.
CC    -!- PTM: Phosphorylated by cyclic GMP-dependent protein kinase (By
CC        similarity).
CC    -!- SIMILARITY: Contains 1 RGS domain.
CC    ------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ------------------------------
DR    EMBL; U27768; AAC50395.1; -.
DR    EMBL; AF493928; AAM12642.1; -.
DR    EMBL; BC000737; AAH00737.1; -.
DR    EMBL; BC051869; AAH51869.1; -.
DR    PIR; S78221; S78221.
DR    HSSP; P49799; 1AGR.
DR    Genew; HGNC:10000; RGS4.
DR    MIM; 602516; -.
DR    GO; GO:0005516; F:calmodulin binding; TAS.
DR    GO; GO:0005096; F:GTPase activator activity; TAS.
DR    GO; GO:0000188; P:inactivation of MAPK; TAS.
DR    GO; GO:0008277; P:regulation of G-protein coupled receptor pr...; TAS.
DR    InterPro; IPR000342; Regl_Gprotein.
DR    Pfam; PF00615; RGS; 1.
DR    PRINTS; PR01301; RGSPROTEIN.
DR    ProDom; PD001580; Regl_Gprotein; 1.
DR    SMART; SM00315; RGS; 1.
```

```
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Lipoprotein; Palmitate;
KW   Phosphorylation.
FT   DOMAIN       62    178       RGS.
FT   LIPID        12     12       S-palmitoyl cysteine (Probable).
FT   LIPID        12     12       S-palmitoyl cysteine (Probable).
FT   LIPID        95     95       S-palmitoyl cysteine.
SQ   SEQUENCE   205 AA;  23255 MW;  7713F1F7496A698B CRC64;

Query Match             33.5%;  Score 408.5;  DB 1;  Length 205;
Best Local Similarity   49.1%;  Pred. No. 3.7e-24;
Matches    82;  Conservative    28;  Mismatches    48;  Indels     9;  Gaps     2;

Qy     41  AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKE-----TRVSPEEAVKWGESFDKLLSHRD  95
               ||:::|||  ||:|           :|   ||          :|   ||::||
Db     16  AKDMKHRLGFLLQK----------SDSCEHNSSHNKDKVVICQRVSQEEVKKWAESLENLISHEC  71

Qy     96  GLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDF  155
               ||      |  |||||||||||||||||||||||||||||||||||||||||||
Db     72  GLaAFKAFLKSEYSEENIDFWISCEEYKKIKSPSKLSPKAKKIYNEFISVQATKEVNLDS  131

Qy    156  HTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLM  202
               ||||| |||||||||||||    ::|||||||||||||||| ||
Db    132  CTREETSRNMLEPTITCFDEAQKKIFNLMERDSYRRFLKSRFYLDLV  178


RESULT 10
RGS4_MOUSE
ID   RGS4_MOUSE     STANDARD;      PRT;    205 AA.
AC   O08899; Q99L30;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Regulator of G-protein signaling 4 (RGS4).
GN   RGS4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Substantia nigra;
RX   MEDLINE=98086343; PubMed=925526;
RA   Nomoto S., Adachi K., Yang L.X., Hirata Y., Muraguchi S., Kiuchi K.;
RT   "Distribution of RGS4 mRNA in mouse brain shown by in situ
RT   hybridization.";
RL   Biochem. Biophys. Res. Commun. 241:281-287(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast tumor;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
```

```
CC       THEIR INACTIVE GDP-BOUND FORM. ACTIVITY ON G(Z)-ALPHA IS INHIBITED
CC       BY PHOSPHORYLATION OF THE G-PROTEIN. ACTIVITY ON G(Z)-ALPHA AND
CC       G(I)-ALPHA-1 IS INHIBITED BY PALMITOYLATION OF THE G-PROTEIN (BY
CC       SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN BRAIN, MODERATELY
CC       LOW LEVELS IN HEART, AND VERY LOW LEVELS IN LUNG, LIVER, AND
CC       SKELETAL MUSCLE.
CC   -!- PTM: Either Cys-2 or Cys-12 or both are palmitoylated (By
CC       similarity).
CC   -!- PTM: Phosphorylated by cyclic GMP-dependent protein kinase (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB004315; BAA20400.1; -.
DR   EMBL; BC003882; AAH03882.1; -.
DR   HSSP; P49799; 1AGR.
DR   MGD; MGI:108409; Rgs4.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Lipoprotein; Palmitate;
KW   Phosphorylation.
FT   DOMAIN       62    178       RGS.
FT   LIPID         2      2       S-palmitoyl cysteine (By similarity).
FT   LIPID        12     12       S-palmitoyl cysteine (By similarity).
FT   LIPID        95     95       S-palmitoyl cysteine (By similarity).
FT   CONFLICT    162    162       R -> K (IN REF. 2).
SQ   SEQUENCE   205 AA;  23288 MW;  5D79581711A1F67C CRC64;

Query Match             33.3%;  Score 405.5;  DB 1;  Length 205;
Best Local Similarity   50.0%;  Pred. No. 6.2e-24;
Matches    81;  Conservative    28;  Mismatches    52;  Indels     1;  Gaps     1;

Qy     41  AKEKRNRLSLLVQKPEFHEDTRS-SRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEA  99
               ||:::|||  ||:|           ::||||||||||||||||||||||||
Db     16  AKDMKHRLGFLLQKSDSCEHSSSHSKKDKVVTCQRVSQEEVKKWAESLENLIIHECGLAA  75

Qy    100  FTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKE  159
               ||||||||||||| ||  |||||||||||||||||||||| ||
Db     76  FKAFLKSEYSENIDFWISCEEYKKIKSPSKLSPKAKKIYNEFISVQATKEVNLDSCTRE  135

Qy    160  VITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDL  201
               ||||| |||||||||||||  ||||||||||||||||||||
Db    136  ETSRNMLQPTITCFDEAQKKIFNLMERDSYRRFLKSRFYLDL  177


RESULT 11
RGS5_MOUSE
ID   RGS5_MOUSE     STANDARD;      PRT;    181 AA.
AC   O08850; Q9D0Z2;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Regulator of G-protein signaling 5 (RGS5).
GN   RGS5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97236828; PubMed=9079700;
```

```
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Reg1_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor.
FT   DOMAIN        64    180       RGS.
FT   CONFLICT      49     49       KA -> NG (IN REF. 1).
FT   CONFLICT      77     77       S -> T (IN REF. 1).
SQ   SEQUENCE   181 AA;   21085 MW;  B4B561CFE3DA9630 CRC64;

Query Match             33.2%;  Score 404.5;  DB 1;  Length 181;
Best Local Similarity   47.3%;
Matches   79;  Conservative   37;  Mismatches   48;  Indels    3;  Gaps    1;

Qy    40 RAKEKRNRLSILLVQKPEFHED--TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG 96
            |||| :|:|:|:||||:    :||::  |:|||:||:|:|::: :|:.|||||:
Db    15 RAKEIKLGILLQKPDSAVDLVIPYNEKPEKPAKAHKPSLEEVLQWRQSLDKLLQNSYG 74

Qy    97 LEAFTRPLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
            :||||||||||:||||||||:|::|   :.||:||:|||:||
Db    75 FASFKSFLKSFSEENLEFWVACENYKKIKSPIKMAEKAKQIYEEFIQTEAPKEVNIDHF 134

Qy   157 TKEVINSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLME 203
            ||::.:::  :::|:|:|||:|:|||||:::|
Db   135 TKDITMKNLVEPSRPSFDLAQKRIYALMEKDSLPRFVRSEFYKELIK 181

RESULT 12
RGS4_RAT
ID   RGS4_RAT     STANDARD;      PRT;   205 AA.
AC   P49799;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Regulator of G-protein signaling 4 (RGS4) (RGP4).
GN   RGS4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96178495; PubMed=8602223;
RA   Druey K.M., Blumer K.J., Kang V.H., Kehrl J.H.;
RT   "Inhibition of G-protein-mediated MAP kinase activation by a new
RT   mammalian gene family.";
RL   Nature 379:742-746(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RA   Zhou M.-Y., Gomez-Sanchez C.E., Gomez-Sanchez E.P.;
RT   "The complete cDNA sequence analysis of the rat RGS4.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 93-159 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=96140645; PubMed=8548815;
RA   Koelle M.R., Horvitz H.R.;
RT   "EGL-10 regulates G protein signaling in the C. elegans nervous
RT   system and shares a conserved domain with many mammalian proteins.";
RL   Cell 84:115-125(1996).
RN   [4]
RP   PHOSPHORYLATION.
RX   MEDLINE=20167219; PubMed=10702309;
RA   Pedram A., Razandi M., Kehrl J., Levin E.R.;
RT   "Natriuretic peptides inhibit G protein activation. Mediation through
RT   cross-talk between cyclic GMP-dependent protein kinase and regulators
RT   of G protein-signaling proteins.";
RL   J. Biol. Chem. 275:7365-7372(2000).
RN   [5]
RP   INHIBITION.
RX   MEDLINE=98016286; PubMed=9353196;
```

```
RA   Chen C., Zheng B., Han J., Lin S.C.;
RT   "Characterization of a novel mammalian RGS protein that binds to
RT   Galpha proteins and inhibits pheromone signaling in yeast.";
RL   J. Biol. Chem. 272:8679-8685(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. BINDS TO G(I)-ALPHA AND G(O)-ALPHA,
CC       BUT NOT TO G(S)-ALPHA.
CC   -!- TISSUE SPECIFICITY: Expressed in heart and muscle.
CC   -!- SIMILARITY: Contains 1 RGS domain.
```
```
DR   EMBL; U67188; AAB50618.1; -.
DR   EMBL; AK004165; BAB23201.1; -.
DR   EMBL; BC037683; AAH37683.1; -.
DR   HSSP; P49799; 1AGR.
DR   MGD; MGI:1098434; Rgs5.
DR   InterPro; IPR000342; Reg1_Gprotein.
```

```
Qy     41  AKEKRNRLSLLVQKPEFHEDTRS-SRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEA  99
               ||| ::||: :||    || | :  :  |   |||   | |:||||| :|:  |
Db     16  AKDMKGHRLGFLLQKSDSCEHSSSHSKKDKVVTCQRVSQEEVKKWAESLENLINHECGLAA  75

Qy    100  FTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKE  159
               ||| | :|:||||| |||| :     ||| :  :|| | : :|   :  |||||| : ||
Db     76  FKAFLKSEYSEENIDFWISCEEYKKIKSPSKLSPKAKKIYNEFISVQATKEVNLDSCCTRE  135

Qy    160  VITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDL  201
               : :  :::  ||   :: ||: | :|:|||| ||| |||:|
Db    136  ETSRNMLEPTITCPDEAQKKIFNLMEKDSYRRFPLKSRFYLDL  177
```

RESULT 13
RGS3_HUMAN
```
ID   RGS3_HUMAN     STANDARD;      PRT;   519 AA.
AC   P49796; Q8TD59; Q8TD68;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Regulator of G-protein signaling 3 (RGS3) (RGP3).
GN   RGS3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=96178495; PubMed=8602223;
RA   Druey K.M., Blumer K.J., Kang V.H., Kehrl J.H.;
RT   "Inhibition of G-protein-mediated MAP kinase activation by a new
RT   mammalian gene family.";
RL   Nature 379:742-746(1996).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Brain;
RA   Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
RT   "cDNA clones of human proteins involved in signal transduction
RT   sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Skuce C.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Inhibits signal transduction by increasing the GTPase
CC       activity of G protein alpha subunits thereby driving them into
CC       their inactive GDP-bound form.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P49796-1; Sequence=Displayed;
CC       Name=2; Synonyms=RGS3T;
CC         IsoId=P49796-2; Sequence=VSP_005662;
CC   -!- PTM: Phosphorylated by cyclic GMP-dependent protein kinase (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   ------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------
DR   EMBL; U27655; AAC50394.1; -.
DR   EMBL; AF493927; AAM12641.1; -.
DR   EMBL; AF493941; AAM12655.1; -.
DR   EMBL; AL162727; CAC78977.1; -.
DR   PIR; S78089; S78089.
DR   HSSP; P49799; 1AGR.
DR   Genew; HGNC:9999; RGS3.
```

```
RA   Tu Y., Wang J., Ross E.M.;
RT   "Inhibition of brain Gz GAP and other RGS proteins by palmitoylation
RT   of G protein alpha subunits.";
RL   Science 278:1132-1135(1997).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF COMPLEX WITH G(I)-ALPHA.
RX   MEDLINE=97262066; PubMed=9108480;
RA   Tesmer J.J.G., Berman D.M., Gilman A.G., Sprang S.R.;
RT   "Structure of RGS4 bound to AlF4-activated G(i alpha1)
RT   stabilization of the transition state for GTP hydrolysis.";
RL   Cell 89:251-261(1997).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. ACTIVITY ON G(Z)-ALPHA IS INHIBITED
CC       BY PHOSPHORYLATION OF THE G-PROTEIN. ACTIVITY ON G(Z)-ALPHA AND
CC       G(I)-ALPHA-1 IS INHIBITED BY PALMITOYLATION OF THE G-PROTEIN.
CC   -!- PTM: Either Cys-2 or Cys-12 or both are palmitoylated (By
CC       similarity).
CC   -!- PTM: Phosphorylated by cyclic GMP-dependent protein kinase.
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   ------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------
DR   EMBL; U27767; AAC52440.1; -.
DR   EMBL; AF117211; AAD12065.1; -.
DR   EMBL; U32327; AAC52367.1; -.
DR   PDB; 1AGR; 16-JUN-97.
DR   PDB; 1EZT; 15-JAN-01.
DR   PDB; 1EZY; 15-JAN-01.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Lipoprotein; Palmitate;
KW   Phosphorylation; 3D-structure.
FT   DOMAIN       62    178       RGS.
FT   LIPID         2      2       S-palmitoyl cysteine (By similarity).
FT   LIPID        12     12       S-palmitoyl cysteine (By similarity).
FT   LIPID        95     95       S-palmitoyl cysteine (By similarity).
FT   HELIX        53     59
FT   HELIX        60     61
FT   TURN         63     68
FT   HELIX        70     82
FT   TURN         83     84
FT   TURN         86     86
FT   HELIX        87    100
FT   TURN        104    106
FT   TURN        107    118
FT   HELIX       119    119
FT   TURN        121    122
FT   TURN        124    125
FT   HELIX       131    140
FT   TURN        141    142
FT   TURN        146    149
FT   HELIX       150    162
FT   HELIX       163    163
FT   TURN        164    168
FT   HELIX       169    170
FT   HELIX       172    175
FT   TURN        176    176
SQ   SEQUENCE   205 AA;  23248 MW;  9647C0EC909D0F6F CRC64;
```

```
Query Match             33.2%;
Best Local Similarity   49.4%;      Score 404.5;  DB 1;  Length 205;
Matches   80;  Conservative   30;  Mismatches   51;  Indels    1;  Gaps    1;
                                    Pred. No. 7.4e-24;
```

DR   MIM; 602189; -.
DR   GO; GO:0005829; C:cytosol; TAS.
DR   GO; GO:0005096; F:GTPase activator activity; TAS.
DR   GO; GO:0000188; P:inactivation of MAPK; TAS.
DR   GO; GO:0008277; P:regulation of G-protein coupled receptor pr. . .; TAS.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Alternative splicing; Phosphorylation.
FT   DOMAIN   394   510   RGS.
FT   VARSPLIC   1   313   Missing (in isoform 2).
FT   /FTId=VSP_005662.
FT   CONFLICT   305   305   K -> R (IN REF. 2).
SQ   SEQUENCE   519 AA; 56601 MW; F1CFE3F27D4673A0 CRC64;

Query Match   33.2%; Score 404; DB 1; Length 519;
Best Local Similarity   44.2%; Pred. No. 2.5e-23;
Matches   80; Conservative   39; Mismatches   60; Indels   2; Gaps   1;

Qy   31 EETSKEAKIR-AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEEAVKWGESFD 88
         :|:: |: || | :.| :||:| : | :.|: ::||||:||||:|:|||||:|||||
Db 337 DEASRKRKSKNLAKDMKNKLGIFRRRNESPGAPPAGKADKMMKSFKPTSEEALKWGESLE 396

Qy   89 KLLSHRDGLEAFTRFLKTEFSSENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAP 148
         |||| |:|:| : :: ||.||:|:||||||| : : ||| ||:||.|:|||:||||
Db 397 KLLVHKYGLAVFQAFLRTEFSSENLEFWLACEDFKKVKSQSKMASKAKKIFAEYIAIQAC 456

Qy 149 KEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQR 208
         |||||| . . ::: ||| ::|: :   ::|:|:|||||:|||||:||||:|||||
Db 457 KEVNLDSYTREHTKDNLQSVTRGCFDLAQKRIFGLMEKDSYPRFLRSDLYLDLINQKKMS 516

Qy 209 P 209
         |
Db 517 P 517

RESULT 14
RGS8_HUMAN
ID   RGS8_HUMAN   STANDARD;   PRT;   180 AA.
AC   P57771;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Regulator of G-protein signaling 8 (RGS8).
GN   RGS8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21218927; PubMed=11318611;
RA   Sood R., Bonner T.I., Makalowska I., Stephan D.A., Robbins C.M.,
RA   Connors T.D., Morgenbesser S.D., Su K., Faruque M.U., Pinkett H.,
RA   Graham C., Baxevanis A.D., Klinger K.W., Landes G.M., Trent J.M.,
RA   Carpten J.D.;
RT   "Cloning and characterization of 13 novel transcripts and the human
RT   RGS8 gene from the 1q25 region encompassing the hereditary prostate
RT   cancer (HPC1) locus.";
RL   Genomics 73:211-222(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
RT   "cDNA clones of human proteins involved in signal transduction
RT   sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC   ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC   THEIR INACTIVE GDP-BOUND FORM. PREFERENTIALLY BINDS TO G(O)-ALPHA
CC   AND G(I)-ALPHA-3 (BY SIMILARITY).

CC   -!- SIMILARITY: Contains 1 RGS domain.
DR   EMBL; AF297015; AAG45337.1; -.
DR   EMBL; AF300649; AAG18443.1; -.
DR   HSSP; P49799; 1AGR.
DR   Genew; HGNC:16810; RGS8.
DR   MIM; 607189; -.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor.
FT   DOMAIN   56   171   RGS.
FT   CONFLICT   1   9   MAALIMPRR -> MWNTLTRSLSDHPVGKDPQAMRTGQRQ
FT   (IN REF. 1).
SQ   SEQUENCE   180 AA; 20917 MW; 198965B4C27F64C9 CRC64;

Query Match   33.0%; Score 402.5; DB 1; Length 180;
Best Local Similarity   46.5%; Pred. No. 9e-24;
Matches   80; Conservative   32; Mismatches   55; Indels   5; Gaps   2;

Qy   40 RAKEKRNRLSLLVQKPEFHEDTRS---SRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG 96
         : | .|:| :||:|:|:||||||| :| : ::.: :. :.:|| ||::|::|::||
Db   9 RNKGMRTRLGCLSHKSDSCSDFTAILPDKPNRALK--RLSTEEATRWADSFDVLLSHKYG 66

Qy   97 LEAFTRFLKTEFSSENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
         : |: : |||||||:|||| |:: :||: : | |::: .||: |:|:| :||||
Db   67 VAAFRAFLKTEFSSENLEFWLACEEFKKTRSTAKLVSKAHRIFEEFVDVQAPREVNIDFQ 126

Qy 157 TKEVITNSITQPTLHSFDAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQR 208
         |:|: |:| : :|| |: .:||:|::|||||:|||:: : |::| :|
Db 127 TREATRKNLQEPSLTCFDQAQGKVHSLMEKDSYPRFLRSKMYLDLLSQSQRR 178

RESULT 15
RGS8_RAT
ID   RGS8_RAT   STANDARD;   PRT;   180 AA.
AC   P49804;
DT   01-OCT-1996 (Rel. 34, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Regulator of G-protein signaling 8 (RGS8).
GN   RGS8.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. AND FUNCTION.
RC   TISSUE=Hippocampus;
RX   MEDLINE=9805413; PubMed=9394004;
RA   Saitoh O., Kubo Y., Miyatani Y., Asano T., Nakata H.;
RT   "RGS8 accelerates G-protein-mediated modulation of K+ currents.";
RL   Nature 390:525-528(1997).
RN   [2]
RP   SEQUENCE OF 87-153 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=9640645; PubMed=8548815;
RA   Koelle M.R., Horvitz H.R.;
RT   "EGL-10 regulates G protein signaling in the C. elegans nervous
RT   system and shares a conserved domain with many mammalian proteins.";
RL   Cell 84:115-125(1996).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE

```
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. PREFERENTIALLY BINDS TO G(O)-ALPHA
CC       AND G(I)-ALPHA-3.
CC   -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN BRAIN. VERY
CC       LITTLE EXPRESSION DETECTED IN OTHER TISSUES.
CC   -!- DEVELOPMENTAL STAGE: DETECTED IN 13-DAY OLD EMBYOS. EXPRESSION
CC       INCREASES GRADUALLY IN LATER EMBRYOS AND MARKEDLY IN NEONATES TO
CC       ADULTS.
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   --------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute.  There are no restrictions on its
CC       use by non-profit institutions as long  as its content is in no  way
CC       modified and this statement is not removed.  Usage by and for commercial
CC       entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC       or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; AB006013; BAA23680.1; -.
DR   EMBL; U32432; AAC52369.1; -.
DR   HSSP; P49799; 1AGR.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor. RGS.
FT   DOMAIN       56    171       RGS.
SQ   SEQUENCE   180 AA;  20949 MW;  00FC35E4C278E857 CRC64;
```

```
Query Match        33.0%;  Score 402.5;  DB 1;  Length 180;
Best Local Similarity  46.5%;  Pred. No. 9e-24;
Matches   80;  Conservative  32;  Mismatches  55;  Indels  5;  Gaps  2;

Qy   40 RAKEKRNRLSLIVQKPEFHEDTRS---SRSGHLAKETRVSPEEAVKWGESPDKLLSHRDG 96
            :||  |  :: |   |      :    | ||| :| :|||| || ||
Db    9 RNKGMRTRLGCLSHKSDSCSDFTAILPDKPNRALK--RLSTEEATRWADSFDVLLSHKYG 66

Qy   97 LEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
          : | ||  |||||:||: | |||   :::| | |:|   || ||:||: |:|:
Db   67 VAAFRAFLKTEFSEENLEFWLACEEFKKTRSTAKLVTKAHRIFEEFVDVQAPREVNIDFQ 126

Qy  157 TKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRPLKSDIYLDLMEGRPQR 208
          || : :  :|| | :|  |  :::|||  ||:|| :|||: :|| :|
Db  127 TREATRKNMQEPSLTCFDQAQGKVHSLMEKDSYPRFLRSKMYLDLLSQSQRR 178
```

Search completed: August 20, 2004, 16:41:29
Job time : 25.5 secs

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      August 20, 2004, 15:52:01 ; Search time 74.5 Seconds
                                        (without alignments)
                                        995.259 Million cell updates/sec

Title:       US-09-894-749-2
Perfect score: 1218
Sequence:    1 METTLLFFSQINMCESKEKT.........SRSFTCNEFQDVQSDVAIWL 235

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:    1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:    1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SPTREMBL_25:*
             1: sp_archea:*
             2: sp_bacteria:*
             3: sp_fungi:*
             4: sp_human:*
             5: sp_invertebrate:*
             6: sp_mammal:*
             7: sp_mhc:*
             8: sp_organelle:*
             9: sp_phage:*
            10: sp_plant:*
            11: sp_rodent:*
            12: sp_virus:*
            13: sp_vertebrate:*
            14: sp_unclassified:*
            15: sp_rvirus:*
            16: sp_bacteriap:*
            17: sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
Result         %     Query
No.  Score  Match Length DB  ID       Description
---- -----  ----- ------ --  ------   -----------
  1  423.5  34.8   213   13  Q7ZZS4   Q7zzs4 gallus gall
  2  423.5  34.8   225   13  Q7ZZS5   Q7zzs5 gallus gall
  3  416    34.2   283   13  Q7SYI2   Q7syi2 gallus gall
  4  416    34.2   408   13  Q7SYI1   Q7syi1 gallus gall
  5  416    34.2   441   13  Q7SYI0   Q7syi0 gallus gall
  6  416    34.2   799   13  Q7SYH9   Q7syh9 gallus gall
  7  413    33.9   182   13  Q7T2D3   Q7t2d3 brachydanio
  8  410.5  33.7   208   13  Q7SZC6   Q7szc6 gallus gall
  9  409    33.6   930   11  Q925G9   Q925g9 mus musculu
 10  408.5  33.5   181    6  Q864Z2   Q864z2 sus scrofa
 11  405.5  33.3   180   11  Q8BXT1   Q8bxt1 mus musculu
 12  404    33.2   917    4  Q8NFN4   Q8nfn4 homo sapien
 13  404    33.2   917    4  Q8WXA0   Q8wxa0 homo sapien
 14  404    33.2   917    4  Q8IUQ1   Q8iuq1 homo sapien
 15  404    33.2  1093    4  Q8NFN5   Q8nfn5 homo sapien
 16  403    33.1   967   11  Q920Q9   Q920q9 rattus norv
```

```
17    402    33.0  192  4  Q8VVE9   Q8vve9 homo sapien
18    401.5  33.0  196  4  Q95K68   Q95k68 macaca fasc
19    401    32.9  192  4  Q8WV02   Q8wv02 homo sapien
20    400    32.8  284  4  Q96NV5   Q96nv5 homo sapien
21    400    32.8  319  4  Q8NFN6   Q8nfn6 homo sapien
22    399.5  32.8  164  6  Q8HXV9   Q8hxv9 ovis aries
23    391    32.1  168  13 Q7SYI3   Q7syi3 gallus gall
24    369    30.3  235  13 Q80V16   Q80v16 mus musculu
25    366.5  30.1  204  13 Q919D9   Q919d9 xenopus lae
26    365.5  30.0  201  11 Q7TNU9   Q7tnu9 mus musculu
27    328    26.9  271  5  Q9V888   Q9v888 drosophila
28    328    26.9  274  5  Q8T017   Q8t017 drosophila
29    324    26.6  208  11 Q8C5J7   Q8c5j7 mus musculu
30    323    26.5  220  13 Q7TQ6    Q7toq6 xenopus lae
31    321    26.4  220  13 Q8QHK0   Q8qhk0 xenopus lae
32    317    26.0  210  11 Q8C5F3   Q8c5f3 mus musculu
33    316    26.0  230  11 Q8BR34   Q8br34 mus musculu
34    315.5  25.9  181  5  Q9TVK0   Q9tvk0 caenorhabdi
35    312.5  25.7  199  13 Q8AYF2   Q8ayf2 gallus gall
36    312    25.6  244  5  Q96842   Q96842 drosophila
37    309.5  25.4  194  11 Q8BFU4   Q8bfu4 mus musculu
38    307    25.2  150  4  Q8TD61   Q8td61 homo sapien
39    306.5  25.2  194  11 Q8CGT5   Q8cgt5 mus musculu
40    286.5  23.5  181  4  Q96GN0   Q96gn0 homo sapien
41    286    23.5  158  11 Q8K443   Q8k443 mus musculu
42    284.5  23.4  475  4  Q7Z4K3   Q7z4k3 homo sapien
43    284.5  23.4  477  4  Q7Z4K4   Q7z4k4 homo sapien
44    284.5  23.4  480  4  Q7Z4K5   Q7z4k5 homo sapien
45    281.5  23.1  424  5  Q95Q48   Q95q48 caenorhabdi
```

ALIGNMENTS

```
RESULT 1
Q7ZZS4
ID   Q7ZZS4        PRELIMINARY;      PRT;   213 AA.
AC   Q7ZZS4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G-protein signaling protein 2.
GN   RGS2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tosetti P., Parente V., Taglietti V., Dunlap K., Toselli M.;
RT   "Concentration-dependent effects of chick RGS2L on neuronal L-type Ca
RT   channel modulation by bradykinin.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF502149; AAP30802.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE 213 AA;  24786 MW;  B0CAD39ACC496D1E CRC64;

Query Match             34.8%;  Score 423.5;  DB 13;  Length 213;
Best Local Similarity   46.9%;  Pred. No. 1.2e-26;
Matches   84; Conservative   34; Mismatches  58; Indels    3; Gaps   1;

Qy    31 EETSKEAKIR---AKEKRNLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEEAVKWGESF 87
              ||||   :: :||:||| |::|||||     ::||::||:|||
Db    27 EEADKGRMKRTIIKDWKKKLSYFLQNSSRSSKVKSTKAGKHHTYFRPSPEEARLWSEAF 86

Qy    88 DKLLSHRDGLEAFTRFLKTFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDA 147
```

```
Db           |:||::: .|| |||::|| |||||||| ||:: |||| ||||
          87 DELLANKYGVAAFRAFLKSEFCEENIEFWLACEDFKKTKSPQKLTLKAKKIYNDFIEKEA 146

Qy       148 PKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRP 206
             ||::::|||| || ::|:| |
Db       147 PKEINIDFQTKNMIAQNLQEATHTCFSAAQKRVYSLMENNSYPRFLESEFYQELCKKPP 205

RESULT 2
Q7ZZS5
ID   Q7ZZS5;
AC   Q7ZZS5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G-protein signaling protein 2 long isoform.
GN   RGS2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tosetti P., Parente V., Taglietti V., Dunlap K., Toselli M.;
RT   "Concentration-dependent effects of chick RGS2L on neuronal L-type Ca
RT   channel modulation by bradykinin.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF502148; AAP30801.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE  225 AA;  25956 MW;  5F8C5FAE2C665E81 CRC64;

Query Match          34.8%;  Score 423.5;  DB 13;  Pred. No. 1.2e-26;
Best Local Similarity  46.9%;
Matches  84;  Conservative  34;  Mismatches  58;  Indels  3;  Gaps  1;

Qy        31 EETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEEAVKWGESF 87
             ||||| ::: :.||
Db        39 EEADKGRMKRTIIKDWKKKLSYFLQNSSRSSKVKSTKAGKHHTYFRPSPEEARLWSEAF 98

Qy        88 DKLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDA 147
             |:||| .::|||
Db        99 DELLANKYGVAAFRAFLKSEFCEENIEFWLACEDFKKTKSPQKLTLKAKKIYNDFIEKEA 158

Qy       148 PKEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRP 206
Db       159 PKEINIDFQTKNMIAQNLQEATHTCFSAAQKRVYSLMENNSYPRFLESEFYQELCKKPP 217

RESULT 3
Q7SYI2
ID   Q7SYI2;
AC   Q7SYI2;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G protein signaling 3 RGS3ss isoform.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=DRG neuron;
RX   MEDLINE=22684492; PubMed=12771384;
RA   Tosetti P., Pathak N., Jacob M.H., Dunlap K.;
RT   "RGS3 mediates a calcium-dependent termination of G protein signaling
RT   in sensory neurons.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7337-7342(2003).
DR   EMBL; AY124774; AAM94020.1; -.
SQ   SEQUENCE  283 AA;  32070 MW;  5C9871357F00A195 CRC64;

Query Match          34.2%;  Score 416;  DB 13;  Length 283;
Best Local Similarity  47.7%;  Pred. No. 6.6e-26;
Matches  83;  Conservative  36;  Mismatches  53;  Indels  2;  Gaps  1;

Qy        31 EETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEEAVKWGESFD 88
             ||||| ::: :.||
Db       101 EETSRKKKRSRSLAKDMKNRLGIFRRRNESPGANPSSKLDKVLKSLKPAPEEALKWGESLE 160

Qy        89 KLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAP 148
             |||:|: ||||
Db       161 KLLLHKYGLAAFRAFLRTEFSEENLEFWLACEBYKKIKSQSKMVSKAKKIFAEYIAIQSC 220

Qy       149 KEVNLDFHTKEVITNSIQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLM 202
Db       221 KEVNLDSYTREHTKENLQNITRGCFDLAQKRIYGLMEKDSYPRFLRSDLYLDII 274

RESULT 4
Q7SYI1
ID   Q7SYI1;
AC   Q7SYI1;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G protein signaling 3 RGS3s isoform.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=DRG neuron;
RX   MEDLINE=22684492; PubMed=12771384;
RA   Tosetti P., Pathak N., Jacob M.H., Dunlap K.;
RT   "RGS3 mediates a calcium-dependent termination of G protein signaling
RT   in sensory neurons.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7337-7342(2003).
DR   EMBL; AY124775; AAM94021.1; -.
SQ   SEQUENCE  408 AA;  45438 MW;  75B50794F22B74B5 CRC64;

Query Match          34.2%;  Score 416;  DB 13;  Length 408;
Best Local Similarity  47.7%;  Pred. No. 1e-25;
Matches  83;  Conservative  36;  Mismatches  53;  Indels  2;  Gaps  1;

Qy        31 EETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEEAVKWGESFD 88
             ||||| ::: :.||
Db       226 EETSRKKKRSRSLAKDMKNRLGIFRRRNESPGANPSSKLDKVLKSLKPAPEEALKWGESLE 285

Qy        89 KLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAP 148
             |||:|: ||||
Db       286 KLLLHKYGLAAFRAFLRTEFSEENLEFWLACEEYKKIKSQSKMVSKAKKIFAEYIAIQSC 345

Qy       149 KEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLM 202
Db       346 KEVNLDSYTREHTKENLQNITRGCFDLAQKRIYGLMEKDSYPRFLRSDLYLDII 399

RESULT 5
Q7SYI0
ID   Q7SYI0;
AC   Q7SYI0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G protein signaling 3 (Fragment).
OS   Gallus gallus (Chicken).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=DRG neuron;
RX   MEDLINE=22684492; PubMed=12771384;
RA   Tosetti P., Pathak N., Jacob M.H., Dunlap K.;
RT   "RGS3 mediates a calcium-dependent termination of G protein signaling
RT   in sensory neurons.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7337-7342(2003).
DR   EMBL; AY124776; AAM94022.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   441 AA;  48612 MW;  F5BB2CC9C25EB70F CRC64;

Query Match             34.2%;  Score 416;  DB 13;  Length 441;
Best Local Similarity   47.7%;  Pred. No. 1.1e-25;
  Matches   83;  Conservative   36;  Mismatches   53;  Indels    2;  Gaps    1;

Qy     31 EETSKEAKIR--AKEKRNRLSLLVQKPFHEDTRSSRSGHLAKETRVSPEEAVKWGESFD 88
          ||||||  || :||||||||||  : |: |||: |  | :||||||||||:|||||||
Db    259 EETSRKKRSRSLAKDMKNRLGIFRRRNESPGANPSSKLDKVLKSLKPAPEBALKWGESLE 318

Qy     89 KLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAP 148
          |||:||  ||||||||||||:|||||||||:||||||:||||||||||||||||||||:|
Db    319 KLLLHKYGLAAFRAFLRTEFSEENLEFWLACEYKKIKSQSKMVSKAKKIFAEYIAIQSC 378

Qy    149 KEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLM 202
          ||||||  || :||: |||| |||| |  |||||||| |||||||:||||||:
Db    379 KEVNLDSYTREHTKENLQNITRGCFDLAQKRIYGLMEKDSYPRFLRSDLYLDII 432


RESULT 6
Q7SYH9
ID   Q7SYH9;
AC   Q7SYH9;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G protein signaling 3 RGS3L isoform.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=DRG neuron;
RX   MEDLINE=22684492; PubMed=12771384;
RA   Tosetti P., Pathak N., Jacob M.H., Dunlap K.;
RT   "RGS3 mediates a calcium-dependent termination of G protein signaling
RT   in sensory neurons.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7337-7342(2003).
DR   EMBL; AY124777; AAM94023.1; -.
SQ   SEQUENCE   799 AA;  89511 MW;  D0A818CF08822B2A CRC64;

Query Match             34.2%;  Score 416;  DB 13;  Length 799;
Best Local Similarity   47.7%;  Pred. No. 2.3e-25;
  Matches   85;  Conservative   36;  Mismatches   53;  Indels    2;  Gaps    1;

Qy     31 EETSKEAKIR--AKEKRNRLSLLVQKPFHEDTRSSRSGHLAKETRVSPEEAVKWGESFD 88
          ||||||  || :||||||||||  : |: |||: |  | :||||||||||:|||||||
Db    617 EETSRKKRSRSLAKDMKNRLGIFRRRNESPGANPSSKLDKVLKSLKPAPEBALKWGESLE 676

Qy     89 KLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAP 148
          |||:||  ||||||||||||:|||||||||:||||||:||||||||||||||||||||:|
Db    677 KLLLHKYGLAAFRAFLRTEFSEENLEFWLACEYKKIKSQSKMVSKAKKIFAEYIAIQSC 736

Qy    149 KEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLM 202
          ||||||  || :||: |||| |||| |  |||||||| |||||||:||||||::
Db    737 KEVNLDSYTREHTKENLQNITRGCFDLAQKRIYGLMEKDSYPRFLRSDLYLDII 790
```

```
RESULT 7
Q7T2D3
ID   Q7T2D3;
AC   Q7T2D3;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., McKernan K.J., Malek J.A., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Peters G.J., Abramson R.D.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC054594; AAH54594.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   182 AA;  21053 MW;  4D502817AA8CE4FF CRC64;

Query Match             33.9%;  Score 413;  DB 13;  Length 182;
Best Local Similarity   49.7%;  Pred. No. 6.8e-26;
  Matches   85;  Conservative   29;  Mismatches   47;  Indels   10;  Gaps    2;

Qy     40 RAKEKRNRLSLLVQKPEF------HEDTRSSRSGHLAKETRVSPEEAVKWGESFDKLLS 92
          |||| ||| :||:||||          ||| : | :||||||||| |: :||||||||
Db     15 RAKEIKTKLGTLLQKPENSIDLIIFYQEKPEKKPBKLQKAT--PEEAAQWRESLDKVLS 71

Qy     93 HRDGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVN 152
          |||||||| :||| | | |:||| :||| ||||:|| |:|||| |  || || |:| ||
Db     72 NSYGLATFKSFLRSEFSEENIEFWEACEDFKKTKNPLKMATKAKKIYEDFIQTGGPKEVN 131

Qy    153 LDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLME 203
          :|| ||:|||: | |: |   |  || :|  | :|| |||:|||||:::::
Db    132 IDHFTKDVTLRNLVDLSSSTFELAQRIYTLMEKDSFGRFLRSDQYQELLK 182


RESULT 8
Q7SZC6
ID   Q7SZC6;
AC   Q7SZC6;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G-protein signaling 4.
OS   Gallus gallus (Chicken).
```

```
Best Local Similarity   45.3%;  Pred. No. 1e-24;
Matches    82;  Conservative   37;  Mismatches   60;  Indels    2;  Gaps    1;

Qy     31 EETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEEAVKWGESFD 88
          :.|.:|.|:|  |||||||||||||||||:||:||||||||||||||||||||||||||
Db    748 DEASRKRKSQNIAKDMKNKLAIFRRRNESPGAQPASKTDKTTKSFKPTSEEALKWSESLE 807

Qy     89 KLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAP 148
          ||||||||| ..|.||||||||||||||||||||||||||||||||||||||||||||
Db    808 KLLLHKYGLEVFQAFLRTEFSEENLEFWIACEDFKKVKSQSKVAAAKAKKIFAEFIAIQAC 867

Qy    149 KEVNLDFHTKEVITNSITQPTLHSFDAAQSRVQLMEQDSYTRFLKSDIYLDLMEGRPQR 208
          ||||||: .|  :|  :||:|||||| |||||||||||||||||||||||||||||||||
Db    868 KEVNLDSYTREHTKENLQSITRGCFDLAQKRIFGLMEKDSYPRFLRSDIYLRSDIYLLINQKMS 927

Qy    209 P 209
Db    928 P 928


RESULT 10
Q86422
ID   Q86422;                      PRELIMINARY;        PRT;    181 AA.
AC   Q86422;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G-protein signalling 5.
GN   RGS5.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Aorta;
RA   Weiss B., Hehlgans S.I., Wietelmann A.B., Zeyer A., Richter M.,
RA   Kloevekorn W.P., Zimmermann R., von der Ahe D.;
RT   "Differential gene expression in atherosclerotic coronary arteries.";
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ549925; CAD71257.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE   181 AA;  21054 MW;  BEB98C7A4A99D669 CRC64;

Query Match             33.5%;  Score 408.5;  DB 6;  Length 181;
Best Local Similarity   47.9%;  Pred. No. 1.6e-25;
Matches    80;  Conservative   36;  Mismatches   48;  Indels    3;  Gaps    1;

Qy     40 RAKEKRNRLSLLVQKPEFHED---TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG 96
          ||||||||:||:||||||   .: :|:|||:||| :||.:|||||||||||||||||||
Db     15 RAKEIKIKLGILLQKPESAVDLVIPYNEKPDKPVKIQKPSLDEALQWRDSLDKLLQNNYG 74

Qy     97 LEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
          |||.|||||:|||||||||||:|||||:|.||:|||||||||||:|:|:|||||||:
Db     75 LASFKSFLKSEFSENLEFWMACEDYKKIKSPVKVAEKAKKIYEFIQSEAPKEVNIDHF 134

Qy    157 TKEVITNSITQPTLHSFDAAQSRVVQLMEQDSYTRFLKSDIYLDLME 203
          |||.:||.:||:|||:|| .:..|:::|||:|:||:|.:|:
Db    135 TKEITMKNLVEPSPSSFDVAQKRIYALMEKDSLPRFVRSEFYQEFIK 181


RESULT 11
Q8BXT1
ID   Q8BXT1;                      PRELIMINARY;        PRT;    180 AA.
AC   Q8BXT1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=D4 hindbrain;
RA   Grillet N., Dubreuil V., Dufour H., Goridis C., Brunet J.-F.;
RT   "Dynamic expression of RGS4 in the developing nervous system and
RT   control by the paired-like homeoprotein Phox2b.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY297457; AAP57222.1; -.
SQ   SEQUENCE   208 AA;  23523 MW;  838B31750FF0BA68 CRC64;

Query Match             33.7%;  Score 410.5;  DB 13;  Length 208;
Best Local Similarity   45.2%;  Pred. No. 1.3e-25;
Matches    85;  Conservative   38;  Mismatches   56;  Indels    9;  Gaps    5;

Qy     41 AKEKRNRLSLLVQKPEFHE-DTRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEA 99
          ||:.|||.||:|||||.:  |||||||||||||||||||:|||||||..||||||||||
Db     16 AKDMKHRLGVLLQKSDSCDYGSSQGKKEKVSSSQRVSQEEVKKWAESLENLIHHDRGLAA 75

Qy    100 FTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKE 159
          |||.|||||||||||||||:.||  :: ||||||||||||||||.|:|:|||||||.:.
Db     76 FRAFLKSEYSEENIEFWVSCEDYKKTKSPAKLSTKARKIYDEFISVQATKEVNLDSCTRE 135

Qy    160 VITNSITQPTLHSFDAAQSRVVQLMEQDSYTRFLKSDIYLDLMEGRPQR----PTNLRR- 214
          :.::.:||:|||.|| :: ||||||:|||||||||||||||||:|..: :  :|  ||:
Db    136 KTSHNMEPTLSCFDEAQRKIFTLMEKDSYRRFLKSP-YLDLVS--PPRAGCGPENCKRA 192

Qy    215 RSRSFTCN 222
          ....:
Db    193 HAHALDCN 200


RESULT 9
Q925G9
ID   Q925G9;                      PRELIMINARY;        PRT;    930 AA.
AC   Q925G9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   PDZ-RGS3 protein.
GN   RGS3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21197945; PubMed=11301003;
RA   Lu Q., Sun E.E., Klein R.S., Flanagan J.G.;
RT   "Ephrin-b reverse signaling is mediated by a novel PDZ-RGS protein and
RT   selectively inhibits G protein-coupled chemoattraction.";
RL   Cell 105:69-79(2001).
CC   -!- SIMILARITY: CONTAINS 1 RGS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
DR   EMBL; AF350047; AAK38878.1; -.
DR   MGD; MGI:1354734; Rgs3.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   InterPro; IPR001478; PDZ.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00595; PDZ; 1.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00228; PDZ; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50106; PDZ; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE   930 AA;  102511 MW;  FB0ECEA164D2F6A1 CRC64;

Query Match             33.6%;  Score 409;  DB 11;  Length 930;
```

DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Regulator of G-protein signaling 8.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Retina;
RX   MEDLINE=22354683; PubMed=12466851;
RA   the FANTOM Consortium;
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573 (2002).
DR   EMBL; AK044337; BAC31874.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE 180 AA; 20963 MW; 00FC35E572785856 CRC64;

Query Match   33.3%; Score 405.5; DB 11; Length 180;
Best Local Similarity   47.1%; Pred. No. 2.7e-25;
Matches   81; Conservative   31; Mismatches   55; Indels   5; Gaps   2;

Qy   40 RAKEKRNRLSLLVQKPEHEDTRS---SRSGHLAKETRVSPEEAVKWGESFDKLLSHRDG 96
         | |||||| || :| ||| |||:   ::|:||||||:||| ||||| ||||||:
Db   9 RNKGMRIRLGCLSHKSDSCSDFTAILPDKPNRALK--RLSTEEATRWAESFDVLLSHKYG 66

Qy   97 LEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFH 156
         |||||||||||||||||| |||||:||:: | :: |||||||||||||:: ::||||
Db   67 VAAFRAFLKTEFSEENLEFWLACEEFKKTRSTAKLVTKAHRIFEEFVDVQAPREVNIDFQ 126

Qy 157 TKEVITNSITQPTLHSPDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQR 208
         |:|: ::: |||||||| ||||||:: ||||| |||||||||:||||||:|
Db 127 TREATRKNMQEPSLTCFDQAQGKVHSLMEKDSYPRFLRSKMYLDLLSQSQRR 178

RESULT 12
Q8NFN4
ID   Q8NFN4   PRELIMINARY;   PRT;   917 AA.
AC   Q8NFN4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   RGS3 isoform PDZ-RGS3.
GN   RGS3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22032988; PubMed=12036301;
RA   Kehrl J.H., Srikumar D., Harrison K., Wilson G.L., Shi C.S.;
RT   "Additional 5' Exons in the RGS3 Locus Generate Multiple mRNA
RT   Transcripts, One of Which Accounts for the Origin of Human PDZ-RGS3.";
RL   Genomics 79:860-868(2002).
DR   EMBL; AF490840; AAM33255.1; -.
CC   -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RGS DOMAIN.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   InterPro; IPR001478; PDZ.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00595; PDZ; 1.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.

DR   SMART; SM00228; PDZ; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50106; PDZ; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE 917 AA; 100975 MW; 63EB882A0F52D87C CRC64;

Query Match   33.2%; Score 404; DB 4; Length 917;
Best Local Similarity   44.2%; Pred. No. 2.5e-24;
Matches   80; Conservative   39; Mismatches   60; Indels   2; Gaps   1;

Qy   31 EETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEEAVKWGESFD   88
         ||:|:||||: |||||||||||||| ||||::::|:||||||||||||| ||||||:
Db 735 DEASRKRKSKNLAKDMKNKLGIFRRRNESPGAPPAGKADKMMKSFKPTSEEALKWGESLE 794

Qy   89 KLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAP 148
         ||| ||||:|||| |||||||||| |||| |||||| || |||||||||||:
Db 795 KLLVHKYGLAVFQAFLRTEFSEENLEFWLACEDFKKVKSQSKMASKAKKIFAEYIAIQAC 854

Qy 149 KEVNLDFHTKEVITNSITQPTLHSPDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQR 208
         ||||| |: || : : |: |||: ::||||||||::||||||||||||:|||||| :
Db 855 KEVNLDSYTREHTKDNLQSVTRGCFDLAQKIFGLMEKDSYPRFLRSDLYLDLINQKKMS 914

Qy 209 P 209
         |
Db 915 P 915

RESULT 13
Q8WXA0
ID   Q8WXA0   PRELIMINARY;   PRT;   917 AA.
AC   Q8WXA0;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   PDZ-RGS3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Chatterjee T.K., Fisher R.A.;
RT   "Human PDZ-RGS3.";
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 RGS DOMAIN.
DR   EMBL; AF463495; AAL68829.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   InterPro; IPR001478; PDZ.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00595; PDZ; 1.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00228; PDZ; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50106; PDZ; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE 917 AA; 100975 MW; 3347CFD383BFAFB9 CRC64;

Query Match   33.2%; Score 404; DB 4; Length 917;
Best Local Similarity   44.2%; Pred. No. 2.5e-24;
Matches   80; Conservative   39; Mismatches   60; Indels   2; Gaps   1;

Qy   31 EETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEEAVKWGESFD   88
         ||:|:||||: |||||||||||||| ||||::::|:||||||||||||| ||||||:
Db 735 DEASRKRKSKNLAKDMKNKLGIFRRRNESPGAPPAGKADKMMKSFKPTSEEALKWGESLE 794

Qy   89 KLLSHRDGLEAFTRFLKTEFSEENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAP 148
         ||| ||||:|||| |||||||||| |||| |||||| || |||||||||||:
Db 795 KLLVHKYGLAVFQAFLRTEFSEENLEFWLACEDFKKVKSQSKMASKAKKIFAEYIAIQAC 854

```
Qy     149  KEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQR 208
            |||||| ||| |:|    ::|    ||  ||| ::| :|| || :|||||:
Db     855  KEVNLDSYTREHTKDNLQSVTRGCFDLAQKRIFGLMEKDSYPRFLRSDLYLDLINQKKMS 914

Qy     209  P 209
            |
Db     915  P 915


RESULT 14
Q8IUQ1
ID   Q8IUQ1;              PRELIMINARY;        PRT;   917 AA.
AC   Q8IUQ1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Regulator of G-protein signalling 3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC042555; AAH42555.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   InterPro; IPR001478; PDZ.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00595; PDZ; 1.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00228; PDZ; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50106; PDZ; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE 917 AA;  100920 MW;  AEBCB3427BC2BC14 CRC64;

Query Match             33.2%;  Score 404;  DB 4;  Length 917;
Best Local Similarity   44.2%;  Pred. No. 2.5e-24;
  Matches    80;  Conservative   39;  Mismatches   60;  Indels    2;  Gaps    1;

Qy      31  EETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEEAVKWGESFD 88
            :| :||| || ||: :| : :: : ::     ||||||
Db     735  DEASRKRKSKNLAKDMKNKLGIFRRRNESPGAPPAGKADKWMKSFKPTSEEALKWGESLE 794

Qy      89  KLLSHRDGLEAFTRFLKTEFSSENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAP 148
            ||| |||||| ||| || |||| |||||| ||||:|||||:||||||||||||||:|||
Db     795  KLLVHKYGLAVFQAFLRTEFSEENLEFWLACEDFKKVKSQSKMASKAKKIFAEYIAIQAC 854

Qy     149  KEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQR 208
            |||||| ||| |:|    ::|    ||  ||| ::| :|| || :|||||:
Db     855  KEVNLDSYTREHTKDNLQSVTRGCFDLAQKRIFGLMEKDSYPRFLRSDLYLDLINQKKMS 914

Qy     209  P 209
            |
Db     915  P 915


RESULT 15
Q8NFN5
ID   Q8NFN5;              PRELIMINARY;        PRT;   1093 AA.
AC   Q8NFN5;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   RGS3 isoform C2PA-RGS3.
GN   RGS3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22032998; PubMed=12036301;
RA   Kehrl J.H., Srikumar D., Harrison K., Wilson G.L., Shi C.S.;
RT   "Additional 5' Exons in the RGS3 Locus Generate Multiple mRNA
RT   Transcripts, One of Which Accounts for the Origin of Human PDZ-RGS3.";
RL   Genomics 79:860-868(2002).
CC   -!- SIMILARITY: CONTAINS 1 RGS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
DR   EMBL; AF490839; AAM33254.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   InterPro; IPR000008; C2.
DR   InterPro; IPR008973; C2_CaLB.
DR   InterPro; IPR001478; PDZ.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00168; C2; 1.
DR   Pfam; PF00595; PDZ; 1.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00239; C2; 1.
DR   SMART; SM00228; PDZ; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50004; C2_DOMAIN_2; 1.
DR   PROSITE; PS50106; PDZ; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE 1093 AA;  120970 MW;  B21187E5643436EB CRC64;

Query Match             33.2%;  Score 404;  DB 4;  Length 1093;
Best Local Similarity   44.2%;  Pred. No. 3.1e-24;
  Matches    80;  Conservative   39;  Mismatches   60;  Indels    2;  Gaps    1;

Qy      31  EETSKEAKIR--AKEKRNRLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEEAVKWGESFD 88
            :| :||| || ||: :| : :: : ::     ||||||
Db     911  DEASRKRKSKNLAKDMKNKLGIFRRRNESPGAPPAGKADKWMKSFKPTSEEALKWGESLE 970

Qy      89  KLLSHRDGLEAFTRFLKTEFSSENIEFWIACEDFKKSKGPQQIHLKAKAIYEKFIQTDAP 148
            ||| |||||| ||| || |||| |||||| ||||:|||||:||||||||||||||:|||
Db     971  KLLVHKYGLAVFQAFLRTEFSEENLEFWLACEDFKKVKSQSKMASKAKKIFAEYIAIQAC 1030

Qy     149  KEVNLDFHTKEVITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQR 208
            |||||| ||| |:|    ::|    ||  ||| ::| :|| || :|||||:
Db     1031 KEVNLDSYTREHTKDNLQSVTRGCFDLAQKRIFGLMEKDSYPRFLRSDLYLDLINQKKMS 1090

Qy     209  P 209
            |
Db     1091 P 1091


Search completed: August 20, 2004, 16:47:00
Job time : 78.5 secs
```

OM nucleic - nucleic search, using sw model

Copyright　GenCore version 5.1.6
Copyright (c) 1993 - 2004　Compugen Ltd.

Run on:　　　August 20, 2004, 03:13:58 ; Search time 4611.24 Seconds
　　　　　　　　(without alignments)
　　　　　　　　10940.929 Million cell updates/sec

Title:　　　　US-09-894-749-3
Perfect score:　1164
Sequence:　　1 ttttgtaagaaaatctga.........gtactagtggatccgagctc 1164

Scoring table:　IDENTITY_NUC
　　　　　　　　Gapop 10.0 , Gapext 1.0

Searched:　　3470272 seqs, 2167151699S residues

Total number of hits satisfying chosen parameters:　　　6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
　　　　　　　　Maximum Match 100%
　　　　　　　　Listing first 45 summaries

Database :　GenEmbl:*
　1: gb_ba:*
　2: gb_htg:*
　3: gb_in:*
　4: gb_om:*
　5: gb_ov:*
　6: gb_pat:*
　7: gb_ph:*
　8: gb_pl:*
　9: gb_pri:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

---

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1164	100.0	1164	6	AR165091	AR165091 Sequence
2	1164	100.0	1164	6	BD265699	BD265699 RGS-conta
3	1164	100.0	1164	6	AR214936	AR214936 Sequence
4	1113.2	95.6	1937	10	AB042807	AB042807 Mus muscu
5	1104.4	94.9	1399	10	AF302685	AF302685 Mus muscu
6	601	51.6	2136	9	AK057114	AK057114 Homo sapi
7	601	51.6	2164	9	BC020632	BC020632 Homo sapi
8	599.6	51.5	1840	6	AX299975	AX299975 Sequence
9	599.6	51.5	2144	9	AX299976	AX299976 Sequence
10	599.6	51.5	2144	9	AF268036	AF268036 Homo sapi
11	597.8	51.4	2138	9	AF076642	AF076642 Homo sapi
12	596	51.2	2217	6	AR165090	AR165090 Sequence
13	596	51.2	2217	6	BD265698	BD265698 RGS-conta
14	596	51.2	2217	6	AR214935	AR214935 Sequence
15	535.4	46.0	218670	10	AC102163	AC102163 Mus muscu
16	352.2	30.3	110000		AC112625_1	Continuation (2 of
17	227.4	19.5	94175	9	AL513175	AL513175 Human DNA
18	227.4	19.5	107919	2	AL391274	AL391274 Homo sapi
19	208.8	17.9	302893	2	AC124876	AC124876 Rattus no
20	192.2	16.5	1486	6	AX299963	AX299963 Sequence
21	177.6	15.3	241	5	AX299968	AX299968 Sequence
22	156.4	13.4	1463	5	AF502148	AF502148 Gallus ga
23	156.4	13.4	1517	5	AF502149	AF502149 Gallus ga
24	156.2	13.4	737	10	AF215668	AF215668 Mus muscu
25	156.2	13.4	1354	10	BC023001	BC023001 Mus muscu
26	156.2	13.4	1364	10	AF432916	AF432916 Mus muscu
27	154.6	13.3	1240	10	MMU67187	U67187 Mus musculu
28	154	13.2	494	4	AF472577	AF472577 Ovis arie
29	152.8	13.1	2765	5	BC054594	BC054594 Danio rer
30	152.4	13.1	636	9	AF493926	AF493926 Homo sapi
31	152.4	13.1	636	9	BT007065	BT007065 Homo sapi
32	152.4	13.1	636	12	BT008081	BT008081 Synthetic
33	152.4	13.1	685	10	AF321837	AF321837 Rattus no
34	152.4	13.1	1345	6	AR270715	AR270715 Sequence
35	152.4	13.1	1345	6	AR380698	AR380698 Sequence
36	152.4	13.1	1345	6	AX663626	AX663626 Sequence
37	152.4	13.1	1345	9	HUMGOS8PPC	L13463 Human helix
38	152.4	13.1	1363	10	AY043246	AY043246 Rattus no
39	152.4	13.1	1377	9	BC007049	BC007049 Homo sapi
40	152.4	13.1	1458	10	BC061969	BC061969 Rattus no
41	152.4	13.1	1629	10	AF279918	AF279918 Rattus no
42	150.2	12.9	408	6	AX710064	AX710064 Sequence
43	144.4	12.4	2342	9	AB066513	AB066513 Macaca fa
44	143.6	12.3	540	6	AX300745	AX300745 Sequence
45	143.6	12.3	543	9	AF300649	AF300649 Homo sapi

ALIGNMENTS

RESULT 1
AR165091
LOCUS　　　　AR165091　　　　　　　　1164 bp　　　DNA　　　linear　　PAT 17-OCT-2001
DEFINITION　Sequence 3 from patent US 6274362.
ACCESSION　AR165091
VERSION　　AR165091.1　GI:16238488
KEYWORDS　.
SOURCE　　Unknown.
　ORGANISM　Unknown.
　　　　　　Unclassified.
REFERENCE　1　(bases 1 to 1164)
　AUTHORS　Hodge,M.R. and Yowe,D.
　TITLE　　RGS-containing molecules and uses thereof
　JOURNAL　Patent: US 6274362-A 3 14-AUG-2001;
FEATURES　　　　　　　　Location/Qualifiers

```
source          1..1164
                /organism="unknown"
                /mol_type="unassigned DNA"
ORIGIN

Query Match       100.0%;  Score 1164;  DB 6;  Length 1164;
Best Local Similarity  100.0%;  Pred. No. 6.8e-204;
Matches 1164;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTCAGGATGTTTTCCT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTCAGGATGTTTTCCT 60

Qy     61 ATGAAATAGCATTCATCTGTGGGAGAGAGAGGAAGAAGAAGAGACTAAGGAAATCTGACATCTGTTGGTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATGAAATAGCATTCATCTGTGGGAGAGAGAGGAAGAAGAAGAGACTAAGGAAATCTGACATCTGTTGGTC 120

Qy    121 ACTGGGACAGAATAATGGATATGTCACTGGTTTTCTTCTCTCAATTAAATATGTGTGAATC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACTGGGACAGAATAATGGATATGTCACTGGTTTTCTTCTCTCAATTAAATATGTGTGAATC 180

Qy    181 AAAAGAGAAAACTTTTTCAAACTAATGCATGGGAGGAAGAAGAAACAAGCATCGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAAAGAGAAAACTTTTTCAAACTAATGCATGGGAGGAAGAAGAAACAAGCATCGA 240

Qy    241 GGCCAAAATCAGAGCGAAAGAAAAAAGGAATAGACTAAGTCTTCTCCTCCTACAGAGGCCTGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGCCAAAATCAGAGCGAAAGAAAAAAGGAATAGACTAAGTCTTCTCCTCCTACAGAGGCCTGA 300

Qy    301 CTTCCATGGAGAGACTCAAGCCAGTAGAATCTGCCCCTTGGCCAAAGAAACAAGAGTCTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTTCCATGGAGAGACTCAAGCCAGTAGAATCTGCCCCTTGGCCAAAGAAACAAGAGTCTC 360

Qy    361 TCCTGAAGCAGCAGTGAATCATTGACAAATTGCTCTCCTCATGAGACGATGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCCTGAAGCAGCAGTGAATCATTGACAAATTGCTCTCCTCATGAGACGATGG 420

Qy    421 AGTGGATCGTCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGTGGATCGTCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTG 480

Qy    481 GGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAGCAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAGCAAA 540

Qy    541 GGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTTCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTTCA 600

Qy    601 TACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCAGCCCACTCTCCACAGTTTGATACGGC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCAGCCCACTCTCCACAGTTTGATACGGC 660

Qy    661 ACAAAGCAGAGTGTACCAGCTCATGAGAACATGACACAGTTATAAACGGCTTTTGAAATCTGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACAAAGCAGAGTGTACCAGCTCATGAGAACATGACACAGTTATAAACGGCTTTTGAAATCTGA 720

Qy    721 GACTTACTTACATTGATAGGAGAAGGAGACCTCAGAGACCAACAAACCTTAGGAGAGACGATC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GACTTACTTACATTGATAGGAGAAGGAGACCTCAGAGACCAACAAACCTTAGGAGAGACGATC 780

Qy    781 ACGATCATTTACTTACAATGATTTCCAAGATGTAAAGTCAGATGTTGCCATTGGTTATG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 ACGATCATTTACTTACAATGATTTCCAAGATGTAAAGTCAGATGTTGCCATTGGTTATG 840

Qy    841 AGTAAAGTCATTTGTCTTCTTTGATAGTGTATATCTAAAATATACTAATA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AGTAAAGTCATTTGTCTTCTTTGATAGTGTATATCTAAAATATACTAATA 900

Qy    901 CTAATGTGTACTTCTAAAATATAGCTTGTTGTATAAGAAGAGATGATTTCATTTTAAAAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CTAATGTGTACTTCTAAAATATAGCTTGTTGTATAAGAAGAGATGATTTCATTTTAAAAT 960
```

(right column)

```
Qy    961 ACACCATGCAAATACATATTAAATGTAAGAACTTTTTATATTATACTAAAATAATATTCATC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ACACCATGCAAATACATATTAAATGTAAGAACTTTTTATATTATACTAAAATAATATTCATC 1020

Qy   1021 ATCTATCTTCCGAAATATTTTATGAAATCATCTATCTGATATTCATTCTCRTATAAAATTCTT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ATCTATCTTCCGAAATATTTTATGAAATCATCTATCTGATATTCATTCTCRTATAAAATTCTT 1080

Qy   1081 TATTTCTACAATAACAGTCAGTAAGCAGAAGCTTTGAAGCCGAATTCCAGCACACTGGCG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TATTTCTACAATAACAGTCAGTAAGCAGAAGCTTTGAAGCCGAATTCCAGCACACTGGCG 1140

Qy   1141 GCCGGTACTAGTGGATCCGAGCTC 1164
          ||||||||||||||||||||||||
Db   1141 GCCGGTACTAGTGGATCCGAGCTC 1164


RESULT 2
BD265699
LOCUS           BD265699                1164 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION      RGS-containing molecules and uses thereof.
ACCESSION       BD265699
VERSION         BD265699.1  GI:33075467
KEYWORDS        JP 2002535979-A/2.
SOURCE          Mus sp.
  ORGANISM      Mus sp.
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE       1  (bases 1 to 1164)
  AUTHORS       Hodge,M.R. and Yowe,D.
  TITLE         RGS-containing molecules and uses thereof
  JOURNAL       Patent: JP 2002535979-A 2 29-OCT-2002;
                MILLENNIUM PHARMACEUTICALS INC
COMMENT         OS    Mus sp. (mouse)
                PN    JP 2002535979-A/2
                PD    29-OCT-2002
                PF    04-FEB-2000 JP 2000597306
                PR    04-FEB-1999 US 09/244314
                PI    MARTIN R HODGE,DAVID YOWE
                PC    C12N15/09,C07K14/47,C07K16/18,C12N5/10,C12P21/02,C12Q1/02,C12Q1/02, PC
                C12Q1/68,
                PC    G01N33/53,G01N33/566,C12N15/00,C12N15/00 CC
                RGS-containing molecules and uses thereof
                FH    Key             Location/Qualifiers
                FT                     (134)..(841)).
FEATURES             Location/Qualifiers
     source          1..1164
                     /organism="Mus sp."
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10095"
ORIGIN

Query Match       100.0%;  Score 1164;  DB 6;  Length 1164;
Best Local Similarity  100.0%;  Pred. No. 6.8e-204;
Matches 1164;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTCAGGATGTTTTCCT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTCAGGATGTTTTCCT 60

Qy     61 ATGAAATAGCATTCATCTGTGGGAGAGAGAGGAAGAAGAAGAGACTAAGGAAATCTGACATCTGTTGGTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATGAAATAGCATTCATCTGTGGGAGAGAGAGGAAGAAGAAGAGACTAAGGAAATCTGACATCTGTTGGTC 120

Qy    121 ACTGGGACAGAATAGTCACTGGTTTTCTTCTCTCAATTAAATATGTGTGAATC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACTGGGACAGAATAGTCACTGGTTTTCTTCTCTCAATTAAATATGTGTGAATC 180

Qy    181 AAAAGAGAAAACTTTTTCAAACTAATGCATGGGAGGAAGAAGAAACAAGCATCGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAAAGAGAAAACTTTTTCAAACTAATGCATGGGAGGAAGAAGAAACAAGCATCGA 240

Qy    241 GGCCAAAATCAGAGCCGAAAGAATAGGAATAGGAAGAAGAACTAAGTCTTCTCCTCCTACAGAGGCCTGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGCCAAAATCAGAGCCGAAAGAATAGGAATAGGAAGAAGAACTAAGTCTTCTCCTCCTACAGAGGCCTGA 300
```

REFERENCE   1   (bases 1 to 1164)
AUTHORS   Hodge,M.R. and Yowe,D.
TITLE   RGS-containing molecules and uses thereof
JOURNAL   Patent: US 6410240-A 3 25-JUN-2002;
FEATURES   Location/Qualifiers
source   1..1164
             /organism="unknown"
             /mol_type="genomic DNA"
ORIGIN

Query Match   100.0%; Score 1164; DB 6; Length 1164;
Best Local Similarity 100.0%;
Matches 1164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTCAGGATGTTTTCCT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTCAGGATGTTTTCCT 60

Qy   61 ATGAAATAGCATTCATCTGTGGGAGAGAAGGACTAAGGAAATCTGACATCTGTTGGTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ATGAAATAGCATTCATCTGTGGGAGAGAAGGACTAAGGAAATCTGACATCTGTTGGTC 120

Qy   121 ACTGGGACAGAATATGGATAATGTCACTGGTTTCTTCTCTCAATTAAATATGTGAATC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ACTGGGACAGAATATGGATAATGTCACTGGTTTCTTCTCTCAATTAAATATGTGAATC 180

Qy   181 AAAAGAGAAAACTTTTTCAAACTAATGCAGTTATAAACGCTTTTGAAATCTGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AAAAGAGAAAACTTTTTCAAACTAATGCAGTTATAAACGCTTTTGAAATCTGA 240

Qy   241 GGCCAAAATCAGAGCGAAAGAAATCAGATAGACTAAGTCTTCCTCACAGAGAGGCCTGA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGCCAAAATCAGAGCGAAAGAAATCAGATAGACTAAGTCTTCCTCACAGAGAGGCCTGA 300

Qy   301 CTTCCATGGAAGAGCCAGTAGATGCTAAAAACTGAATTCAGTGAGGAGAAACAAGAGTCTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTTCCATGGAAGAGCCAGTAGATGCTAAAAACTGAATTCAGTGAGGAGAAACAAGAGTCTC 360

Qy   361 TCCTGAAGAAGCAGTGAAATGGCTGAATCATTTGACAAATTGCTTCTCATAGAGATGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCCTGAAGAAGCAGTGAAATGGCTGAATCATTTGACAAATTGCTTCTCATAGAGATGG 420

Qy   421 AGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG 480

Qy   481 GGTCGCCTGTGAAGACTTCAAGAAATGCAAGAACCTCAACAAATCATCCTAAAAGCAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GGTCGCCTGTGAAGACTTCAAGAAATGCAAGAACCTCAACAAATCATCCTAAAAGCAAA 540

Qy   541 GGCAATCTATGAGAAATTCATTCAGAATGATTTCCAAGATGTAAACTCAGATGTTGGCCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GGCAATCTATGAGAAATTCATTCAGAATGATTTCCAAGATGTAAACTCAGATGTTGGCCA 600

Qy   601 TACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGTTTTGATACGGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGTTTTGATACGGC 660

Qy   661 ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTGAAATCTGA 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTGAAATCTGA 720

Qy   721 GACCTACTTACATTGATAGTGTATGTGTATATCTAAAATATACTAATA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GACCTACTTACATTGATAGTGTATGTGTATATCTAAAATATACTAATA 780

Qy   781 ACGATCATTTACTTACAATGATTTCCAAGATGTAAACTCAGATGTTGGTTATG 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ACGATCATTTACTTACAATGATTTCCAAGATGTAAACTCAGATGTTGGTTATG 840

Qy   841 AGTAAAAGTCATTTGTCTTCTTTTGATAGTGTATGTGTATATCTAAAATATACTAATA 900

Db   241 GGCCAAAATCAGAGCGAAAGAAATCAGATAGACTAAGTCTTCCTCACAGAGAGGCCTGA 300

Qy   301 CTTCCATGGAAGAGCCAGTAGATGCTAAAAACTGAATTCAGTGAGGAGAACAAGAGTCTC 360
Db   301 CTTCCATGGAAGAGCCAGTAGATGCTAAAAACTGAATTCAGTGAGGAGAACAAGAGTCTC 360

Qy   361 TCCTGAAGAAGCAGTGAAATGGCTGAATCATTTGACAAATTGCTTCTCATAGAGATGG 420
Db   361 TCCTGAAGAAGCAGTGAAATGGCTGAATCATTTGACAAATTGCTTCTCATAGAGATGG 420

Qy   421 AGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG 480
Db   421 AGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG 480

Qy   481 GGTCGCCTGTGAAGACTTCAAGAAATGCAAGAACCTCAACAAATCATCCTAAAAGCAAA 540
Db   481 GGTCGCCTGTGAAGACTTCAAGAAATGCAAGAACCTCAACAAATCATCCTAAAAGCAAA 540

Qy   541 GGCAATCTATGAGAAATTCATTCAGAATGATTTCATTCTCAATTCTAATAAAATTCTT 600
Db   541 GGCAATCTATGAGAAATTCATTCAGAATGATTTCATTCTCAATTCTAATAAAATTCTT 600

Qy   601 TACTTTCTCACAATAACAGTCAGTAAGAAGAGAAGCTTTGAAGCCGAATTCCAGCACACTGGCG 660
Db   601 TATTTCTCACAATAACAGTCAGTAAGAAGAGAAGCTTTGAAGCCGAATTCCAGCACACTGGCG 660

Qy   1141 GCCGGTACTAGTGGATCCGAGCTC 1164
Db   1141 GCCGGTACTAGTGGATCCGAGCTC 1164

RESULT 3
AR214936
LOCUS   AR214936   1164 bp   DNA   linear   PAT 25-SEP-2002
DEFINITION Sequence 3 from patent US 6410240.
ACCESSION   AR214936
VERSION   AR214936.1   GI:23312889
KEYWORDS
SOURCE   Unknown.
ORGANISM   Unknown.

```
Db   841 AGTAAAAGTCATTTGTCTTCTTTTGATAGTGTATGTGTATATCTAAAATATACTAATA 900

Matches 1115;  Conservative 0;  Mismatches 3;  Indels 0;  Gaps 0;

Qy     1 TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGATGAGCGCTTTATTCAGGATGTTTCCT 60
         |||||||||||||||||||||||||||||||||||||||||  |||||||||||||||
Db    63 TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGATGAGCGCTTAATTCAGGATGTTTCCT 122

Qy    61 ATGAAATAGCATTCATCTGTCGGAGAGAGAGGACTAAGGAAATCTGACATCTGTTGGTC 120
         ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123 ATGAAACAGCATTCATCTGTCGGAGAGAGAGGACTAAGGAAATCTGACATCTGTTGGTC 182

Qy   121 ACTGGGACAGAATATGGATATGTCACTGGTTTCCTCTCTCAATAAATATGTGTGAATC 180
         |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||
Db   183 ACTGGGACAGAATATGGATATGTCACTGGTTTCTCTCTCTCAATAAATATGTGTGAATC 242

Qy   181 AAAAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGGAAAGAGAAGAAACAAGCATCGA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   243 AAAAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGGAAAGAGAAGAAACAAGCATCGA 302

Qy   241 GGCCAAAATCAGAGGCAAAGAAAGAATAGACTAAGTCTTCTCCTACAGAGAGGCCTGA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   303 GGCCAAAATCAGAGGCAAAGAAAGAATAGACTAAGTCTTCTCCTACAGAGAGGCCTGA 362

Qy   301 CTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCTC 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   363 CTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCTC 422

Qy   361 TCCTGAAGAAGCAGTGAAATGGGCTGAAATCATTGACAAATGCTCTCCTCATAGAGATGG 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   423 TCCTGAAGAAGCAGTGAAATGGGCTGAAATCATTGACAAATGCTCTCCTCATAGAGATGG 482

Qy   421 AGTGGATGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   483 AGTGGATGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG 542

Qy   481 GGTCGCCTGTGAAGACTTCAAGAAATGCAAGGACCTCAACAAATCATCCTAAAAGCAA 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   543 GGTCGCCTGTGAAGACTTCAAGAAATGCAAGGACCTCAACAAATCATCCTAAAAGCAA 602

Qy   541 GGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTTCA 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   603 GGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTTCA 662

Qy   601 TACTAAAGAGTAATTGCTAAGAGCATCGCCAGCCAGTCTCCACAGTTTTGATACGGC 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   663 TACTAAAGAGTAATTGCTAAGAGCATCGCCAGCCAGTCTCCACAGTTTTGATACGGC 722

Qy   661 ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTGAAATCGA 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   723 ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTGAAATCGA 782

Qy   721 GACCTACTTACATTGATAGAAGGAAGACCTCAGAGACCAACAAACCTTAGGAGGACGATC 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   783 GACCTACTTACATTGATAGAAGGAAGACCTCAGAGACCAACAAACCTTAGGAGGACGATC 842

Qy   781 ACGATCATTACTTACTTACCAAGATGATTCCAAGATGTTGCCATTGGTATG 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   843 ACGATCATTACTTACTTACCAAGATGATTCCAAGATGTTGCCATTGGTATG 902

Qy   841 AGTAAAAGTCATTTGTCTTCTTTTGATAGTGTATGTGTATCTAAAATATACTAATA 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   903 AGTAAAAGTCATTTGTCTTCTTTTGATAGTGTATGTGTATCTAAAATATACTAATA 962

Qy   901 CTAATGTGTACTTCTCTAAAATATAGCTTGTGTATAAGAAGAGATGATTCATTTTTAAAAT 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   963 CTAATGTGTACTTCTCTAAAATATAGCTTGTGTATAAGAAGAGATGATTCATTTTTAAAAT 1022

Qy   961 ACACCATGCAAATACATATTAAATAAATGTAAGAACTTTTATATTATAAAATAATTCATC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1023 ACACCATGCAAATACATATTAAATAAATGTAAGAACTTTTATATTATAAAATAATTCATC 1082

Qy  1021 ATCTATCTTCCGAAATATTTTATGAAAATCTATCTGATATTCTTATTCTAAAAAATTCTT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1083 ATCTATCTTCCGAAATATTTTATGAAAATCTATCTGATATTCTTATTCTAAAAAATTCTT 1142

Qy   901 CTAATGTGTACTTCTCTAAAATATAGCTTGTGTATAAGAAGAGATGATTCATTTTTAAAAT 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CTAATGTGTACTTCTCTAAAATATAGCTTGTGTATAAGAAGAGATGATTCATTTTTAAAAT 960

Qy   961 ACACCATGCAAATACATATTAAATAAATGTAAGAACTTTTATATTATAAAATAATTCATC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 ACACCATGCAAATACATATTAAATAAATGTAAGAACTTTTATATTATAAAATAATTCATC 1020

Qy  1021 ATCTATCTTCCGAAATATTTTATGAAAATCTATCTGATATTCTTATTCTAAAAAATTCTT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ATCTATCTTCCGAAATATTTTATGAAAATCTATCTGATATTCTTATTCTAAAAAATTCTT 1080

Qy  1081 TATTTCTACAATAACAGTCAGTAAGAACAGAAGCTTGAAGCCGAATCCAGCACACTGGCG 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TATTTCTACAATAACAGTCAGTAAGAACAGAAGCTTGAAGCCGAATCCAGCACACTGGCG 1140

Qy  1141 GCCGGTACTAGTGGATCCGAGCTC 1164
         ||||||||||||||||||||||||
Db  1141 GCCGGTACTAGTGGATCCGAGCTC 1164
```

RESULT 4
AB042807

LOCUS   AB042807   1937 bp   mRNA   linear   ROD 12-JUN-2001
DEFINITION Mus musculus RGS mRNA, complete cds.
ACCESSION   AB042807
VERSION   AB042807.1 GI:14349231
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (sites)
  AUTHORS   Nagata,Y., Oda,M., Nakata,H., Shozaki,Y., Kozasa,T. and Todokoro,K.
  TITLE   A novel regulator of G-protein signaling bearing GAP activity for
            Galphai and Galphaq in megakaryocytes
  JOURNAL   Blood 97 (10), 3051-3060 (2001)
  MEDLINE   21240406
   PUBMED   11342430
REFERENCE   2 (bases 1 to 1937)
  AUTHORS   Todokoro,k., Nagata,Y. and Oda,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (17-MAY-2000) kazuo Todokoro, RIKEN Tsukuba Institute,
            Molecular Cell Science Laboratory; Koyadai, Tsukuba, Ibaraki
            305-0074, Japan (E-mail:todokoro@rtc.riken.go.jp,
            Tel:81-298-36-9075, Fax:81-298-36-9090)
FEATURES   Location/Qualifiers
     source   1..1937
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="BDF-1"
                     /db_xref="taxon:10090"
                     /sex="female"
                     /cell_type="megakaryocyte"
                     /dev_stage="adult"
     gene   1..1937
                     /gene="RGS"
     CDS   196..903
                     /gene="RGS"
                     /codon_start=1
                     /protein_id="BAB60723.1"
                     /db_xref="GI:14349232"
                     /translation="MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEK
                     RNRLSLLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTR
                     FLKTEFSEENIEFWVACEDFKKCKEPQQILKAKAIYEKFIQNDAPKEVNIDFHTKEV
                     IAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIGRPQRPTNLRRSRSRS
                     FTYNDFQDVKSDVAIWL"
ORIGIN

Query Match   95.6%; Score 1113.2; DB 10; Length 1937;
Best Local Similarity 99.7%; Pred. No. 1.3e-194;

```
Qy    301  CTTCCATGCTGAGAGACTCAAGCCAGTAGATCTGCCCCTTGGCCAAAGAAACAAGAGTCTC  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    354  CTTCCATGCTGAGAGACTCAAGCCAGTAGATCTGCCCCTTGGCCAAAGAAACAAGAGTCTC  413

Qy    361  TCCTGAAGAAGCAGTGAAATGGCTGAATCATTGGACAAATTGCTCCTCATAGAGATGG  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    414  TCCTGAAGAAGCAGTGAAATGGCTGAATCATTGGACAAATTGCTCCTCATAGAGATGG  473

Qy    421  AGTGGATGCTTTTACCAGATTCTTAAAACTGAATCAGTGAGAGACATTGAATTTTG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    474  AGTGGATGCTTTTACCAGATTCTTAAAACTGAATCAGTGAGAGACATTGAATTTTG  533

Qy    481  GGTCGCCTGTGAAGACTTCAAGAAATGCAAGAACCTCAACAAATCATCCTAAAGCAAA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    534  GGTCGCCTGTGAAGACTTCAAGAAATGCAAGAACCTCAACAAATCATCCTAAAGCAAA  593

Qy    541  GGCAATCTATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTTCA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    594  GGCAATCTATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTTCA  653

Qy    601  TACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGATACGGC  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    654  TACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGATACGGC  713

Qy    661  ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTGAAATCTGA  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    714  ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTGAAATCTGA  773

Qy    721  GACCTACTTACATTGATGAGAGGAAGACCTCAGAGACCAACAAACCTTAGGAGAGACGATC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    774  GACCTACTTACATTGATGAGAGGAAGACCTCAGAGACCAACAAACCTTAGGAGAGACGATC  833

Qy    781  ACGATCATTTACTTACTACAATGATTTCCAAGATGTAAAGTCAGATGTTGCCATTGGTTATG  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    834  ACGATCATTTACTTACTACAATGATTTCCAAGATGTAAAGTCAGATGTTGCCATTGGTTATG  893

Qy    841  AGTAAAAGTCATTCAGTTGTCTCTTTTGATAGTTGTGTATGCTAAAATATATACTAATA  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    894  AGTAAAAGTCATTCAGTTGTCTCTTTTGATAGTTGTGTATGCTAAAATATATACTAATA  953

Qy    901  CTAATGTGTACTTCTAAAATACATTAAATGTAAGAACTTTTATATATACTAAAATAATTCATC  960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    954  CTAATGTGTACTTCTAAAATACATTAAATGTAAGAACTTTTATATATACTAAAATAATTCATC 1013

Qy    961  ACACCATGCAAATACATTAAATGTAAGAACTTTTTATATTAATTATACTAAAATAATTCATC 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1014  ACACCATGCAAATACATTAAATGTAAGAACTTTTTATATTAATTATACTAAAATAATTCATC 1073

Qy   1021  ATCTATCTTCCGAAATATTTATGAAAATCTATCTGATATTCATTCTAATAAAATTCTT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1074  ATCTATCTTCCGAAATATTTATGAAAATCTATCTGATATTCATTCTAATAAAAATTCTT 1133

Qy   1081  TATTTCTACAATAACAGTCAGTAAGAAGAAGCTTTGAA 1118
           |||||||||||||||||||||||||||||||||||||
Db   1134  TATTTCTACAATAACAGTCAGTAAGAAGAAGCTTTGTA 1171
```

```
Qy   1081  TATTTCTACAATAACAGTCAGTAAGAAGAAGCTTTGAA 1118
           |||||||||||||||||||||||||||||||||||||
Db   1143  TATTTCTACAATAACAGTCAGTAAGAAGAAGCTTTGTA 1180
```

RESULT 5
AF302685

```
LOCUS       AF302685                1399 bp    mRNA    linear   ROD 24-JAN-2001
DEFINITION  Mus musculus regulator of G-protein signaling RGS18 mRNA, complete
            cds.
ACCESSION   AF302685
VERSION     AF302685.1  GI:12407842
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1399)
  AUTHORS   Park,I.K., Klug,C.A., Li,K., Jerabek,L., Li,L., Nanamori,M.,
            Neubig,R.R., Hood,L., Weissman,I.L. and Clarke,M.F.
  TITLE     Molecular cloning and characterization of a novel regulator of
            G-protein signaling from mouse hematopoietic stem cells
  JOURNAL   J. Biol. Chem. 276 (2), 915-923 (2001)
  MEDLINE   21125614
   PUBMED   11042171
REFERENCE   2  (bases 1 to 1399)
  AUTHORS   Park,I.K., Klug,C.A., Li,K., Jerabek,L., Li,L., Nanamori,M.,
            Neubig,R.R., Hood,L., Weissman,I.L. and Clarke,M.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-SEP-2000) Internal Medicine, University of Michigan,
            1500 E. Medical Center Dr., Ann Arbor, MI 48109, USA
FEATURES             Location/Qualifiers
     source          1..1399
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="BA"
                     /db_xref="taxon:10090"
     CDS             187..894
                     /codon_start=1
                     /product="regulator of G-protein signaling RGS18"
                     /protein_id="AAG3657.1"
                     /db_xref="GI:12407843"
                     /translation="MDMSLVFPSQLNMCESKEKTFKLMHGSGKETSIPAKIRAKEK
                     RNRLSLLQRPDFHGEQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTR
                     FLKTEFSEENIEFWVACEDFKKCKEPQOIILKAKAIYEKFIQNDAPKEVNIDFHTKEV
                     IAKSIAQPTLHSFDTAQSRVTQLMEHDSYKRPLKSETYLHLIEGRPQRPTNLRRRSRS
                     FTYNDFQDVKSDVAIWL"
ORIGIN
```

```
Query Match             94.9%;  Score 1104.4;  DB 10;  Length 1399;
Best Local Similarity   99.8%;  Pred. No. 5.8e-193;
Matches 1116;  Conservative  0;  Mismatches  1;  Indels  1;  Gaps  1;
```

```
Qy      1  TTTTTGTAAGAAAATCTGAGGAGAGAGATTCGGGAGAGAAGATCCTTTATTCAGGATGTTTCCT  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     55  TTTTTGTAAGAAAATCTGAGGAGAGAGATTCGGGAGAGAAGATCCTTTATTCAGGATGTTTCCT 114

Qy     61  ATGAAAATAGCCATTCATCCTGTGGGAGGAGAAGGACTAAGGAAATCTGACATCTGTTGGTC 120
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    115  ATGAAAATAGCCATTCATCCTG-GGGAGGAGAAGGACTAAGGAAATCTGACATCTGTTGGTC 173

Qy    121  ACTGGGACAGAATGGGAATATGTCACTGGTTTCTTCTCTCAATTAAATATGTGTGAATC 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    174  ACTGGGACAGAATGGGAATATGGATATGTCACTGGTTTCTTCTCTCAATTAAATATGTGTGAATC 233

Qy    181  AAAAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGGAAAGAAGAACAAGCATCGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    234  AAAAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGGAAAGAAGAACAAGCATCGA 293

Qy    241  GGCCAAAATCAGAGCGAAAGAAAAAGGAATAGCAAAAAGGAATAGACTAAGCTTCCTCCTACAGAGGCCTGA 300
           |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    294  GGCCAAAATCAGAGCGAAAGAAGAAGGAATAGCAAAAAGGAATAGACTAAGCTTCCTCCTACAGAGGCCTGA 353
```

RESULT 6
AK057114

```
LOCUS       AK057114                2136 bp    mRNA    linear   PRI 01-AUG-2002
DEFINITION  Homo sapiens cDNA FLJ32552 fis, clone SPLEN1000089, highly similar
            to Homo sapiens regulator of G-protein signaling 13 mRNA.
ACCESSION   AK057114
VERSION     AK057114.1  GI:16552701
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J.,
```

```
            Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K.,
            Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H.,
            Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K.,
            Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2136)
  AUTHORS   Isogai,T., Otsuki,T. and Sugiyama,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
  COMMENT   NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB) (cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: RAB and
            HRI.
FEATURES             Location/Qualifiers
     source          1..2136
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="SPLEN1000089"
                     /tissue_type="spleen"
                     /clone_lib="SPLEN1"
                     /note="cloning vector: pME18SFL3"
ORIGIN

Query Match     51.6%;  Score 601;  DB 9;  Length 2136;
Best Local Similarity   75.2%;  Pred. No. 1.4e-100;
Matches  816;  Conservative    0;  Mismatches 235;  Indels 34;  Gaps    4;

Qy       1 TTTTTGTAAGAAAAATCTGAGGAAGAAGATTCGGGATAGGCCTTTATTC-AGGATGTTTCC 59
           ||| |||||||| |||| || | |||||||||||||||||||||||    || ||||||||
Db      38 TTCTTTTTGTAAACATTAGACATTACGTGTGATAACTTTTATTCTACTACTACTATATG 97

Qy      60 TATGAAATAGCATTCATCTGTGGGAGAGAGAAGGACTAAGAAAATCTGACATCTGTTGGT 119
           ||||| ||||||| |||||||||||| |||||||| |||||||||||||||||||||
Db      98 TATGGAATAGTATTAATAAATGACTACGGACGATGGTAATAATTAGACATCTTCAT 157

Qy     120 CACTCGGGACAGAGATATGGATATGTCACCGGTTTTCTCTCCAATTAAATAAATATGTGAAT 179
           ||||| ||||||| ||||||||| |||||| |||| |||||||||||||||||||||
Db     158 T--TTAGAGAGAAGATGAAACAACATTGCTTTCTTTTCTCAAATAAATATGTGAAT 215

Qy     180 CAAAAGAGAAAAACTTTTTTCAAACTAATGCATGGGTCAGGGAAAGAAAGAAACAAGCATCG 239
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     216 CAAAAGAAAAAAACTTTTTTCAAGTTAATACATACATGGTTCAGGAGTTCAGGAAAGAAAGAAAGCAAAG 275

Qy     240 AGGCCAAAATCAGAGCGAAAGCCGAAAGAGAATAGACTAAGGTCCTTCCCCTACAGAGGCCTG 299
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     276 AAGCCAAAATCAGAGCGAAAGCCGAAAGAGAATAGACTAAGGTCCTTCCTGTGCCAGAACCTG 335

Qy     300 ACTTCCATGGAAGACACCCGCTCCAGTAGATCTGCCCTCTGCCAAAGAAACAAGAGTCT 359
           ||||||||||||| |||||||||||||||||| |||||||||||||||||||||||||||
Db     336 AGTTTCATGAAGACATGAAAAGCAGGCACTTGGGCACTTGCCAAAGAAACAAGAGTCT 395

Qy     360 CTCCTGAAGAAGCAGTGAAATGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGATG 419
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     396 CCCCTGAAGAAGCAGTGAAATGGGTGAATCATTCATTGACAAACTGCTTTCCCATAGAGATG 455

Qy     420 GAGTGGATGCTTTTACCAGATTCTTAAAAACTGAATTCAGTGAGAGAACATTGAATTTT 479
           ||| ||||||| ||||||||||||| |||||| ||||||||||||||||||||||
Db     456 GACTAGAGGCTTTTACCAGATTTCTTAAAAACTGAATTCAGTGAAGAAATATTGAATTTT 515

Qy     480 GGGGTCGCCTGTGAAGACTTCGAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAA 539
           ||| |||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     516 GGATAGCCTGTGAAGAGTTCAGAAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAAAGCAA 575

Qy     540 AGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTTC 599
```

```
Db     576 AAGCAATATATGAGAAATTTATACAGACTGATGCGCCCAAAGGCCCACTCTCCACAGTTTGATTTTC 635
           |||||||||||||||||| || |||| ||||| ||| ||||||  ||||||||||||||||||||
Qy     600 ATACTAAAGAGTAATTGCTAAGAGCATCGCCCAGCCCACTCCCACCTCCACAGTTTGATACGG 659
           | ||||||||||||||||| || |||||||| ||||||||||||||||||||||||||||||
Db     636 ACACAAAAGAGTCATTACAACACAGCATCACTCAACCTCACCTCCCACAGTTTGATGCTG 695
           ||||||||||| |||| ||||||||||||| ||||||||||||||||||||||||||
Qy     660 CACAAAGCAGAGTGTACCAGTCATGGAACATGACAGTTATAAAGCTTTTGAAATCTG 719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     696 CACAAAGCAGAGTGATCAGTCATGGAACAAGAAGTTATACACGTTTTCTGAAATCTG 755
           |||||||||||| |||||||||||||||| || ||||||| ||||  ||||||||
Qy     720 AGACCTACTTACATTTGATAGAAGAAGACCTCAGAGACCAACAAACCTTAGGACGAT 779
           ||| |||||| |||||||||| |||||||| |||||||||| ||||| ||||| || ||
Db     756 ACATCTATTAGACTTGATGGGAAGGAAGGAAGAAGACCTCAGAGACCAACAAATCTTAGGACGACGAT 815
           ||| ||  ||  ||||||| ||||| || ||||||||||||||||||||||| |||
Qy     780 CACGATCATTCACTTACAATGATTCCAAGATGTAAGTCAGGATGTTGGCATTGGTTAT 839
           |||| |||| ||  ||||||||  ||| || | |||||| || |||||| ||||||
Db     816 CACGCCATTTACCTGCAATGAATTCCAAGATGTACAATCAGATGTTGCCATTGGTTAT 875
           ||| |||| ||  |||||||  |||| |||||| | |||| | ||||| |||||||
Qy     840 GAGTAAAGTCATTGTCTTCTTTGCTCATTTTTATGACAAACTTATACATC------- 899
           |||||||||||| ||  |||||||||||||||||||||||||| ||||
Db     876 AAAGAAAAATGATTTTGCTCATTTTGTGTGTATAAGAGAGATGATTCATTTT----- 919
           ||||| ||||  ||||| ||| |||||  || |||||| |||||||||
Qy     900 ACTAATGTGTACTTCTAAAATATAGCTTGTGTATAAGAGAAGATGATTCATTTT----- 954
           |||  |||||||  |||||| |||||| |||  ||||||||| |||||||
Db     920 --------TGCTTCTAACATCGCATGTTTATGTTAAGATTTGGTCCCATCCTTTAAA 970
           | |||| |||||||| ||||||||| |||||| ||||| ||||||||||
Qy     955 -TAAATACACCATGCAAATACATATTAAATGTGTAAGAACTTTTATTATATACTAAAATA 1013
           ||||| |||||||||||||||||| |||||||||||| |||||||||||||||||
Db     971 CTGAAATATGTCATGTGAAATTATTTTAAAATGTAAAAATGTAAAAACAAAACTTTCTGCTAACAAA 1030
           ||  ||||||||||||  |||||||| ||||||||| ||||| |||||||||||||
Qy    1014 ATTCATCATCTATCTTCCGAAAATATTTTATGAAAATCATCTGATAATTCTATTCTAATAA 1073
           |||||||||||||||||||| |||||||||||||||||  ||| ||||||||||||
Db    1031 ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCATTTCATTTGATGTCATTCATTTA 1090

Qy    1074 AATTC 1078
           |||||
Db    1091 TAATC 1095
```

```
RESULT 7
BC020632
LOCUS       BC020632               2164 bp    mRNA    linear    PRI 04-OCT-2003
DEFINITION  Homo sapiens regulator of G-protein signalling 18, mRNA (cDNA clone
            MGC:22266 IMAGE:4714909), complete cds.
ACCESSION   BC020632
VERSION     BC020632.1  GI:18088322
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2164)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
```

```
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE    22388257
PUBMED     12477932
REFERENCE  2  (bases 1 to 2164)
AUTHORS    Strausberg,R.
TITLE      Direct Submission
JOURNAL    Submitted (03-JAN-2002) National Institutes of Health, Mammalian
           Gene Collection (MGC), Cancer Genomics Office, National Cancer
           Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
           USA
REMARK     NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT    Contact: MGC help desk
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: CLONTECH
           cDNA Library Preparation: CLONTECH Laboratories, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Sequencing Group at the Stanford Human Genome
           Center, Stanford University School of Medicine, Stanford, CA 94305
           Web site:    http://www-shgc.stanford.edu
           Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
           Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
           R. M.

           Clone distribution: MGC clone distribution information can be found
           through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
           Series: IRAL Plate: 37 Row: d Column: 18
           This clone was selected for full length sequencing because it
           passed the following selection criteria: matched mRNA gi: 18640749.
FEATURES             Location/Qualifiers
     source          1..2164
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:22266 IMAGE:4714909"
                     /tissue_type="liver"
                     /clone_lib="NIH_MGC_76"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
     gene            1..2164
                     /gene="RGS18"
                     /note="synonym: RGS13"
                     /db_xref="LocusID:64407"
                     /db_xref="MIM:607192"
     CDS             169..876
                     /codon_start=1
                     /product="regulator of G-protein signalling 18"
                     /protein_id="AAH20632.1"
                     /db_xref="GI:18088323"
                     /db_xref="LocusID:64407"
                     /translation="METTLLFFSQINMCESKEKTFKLIHGSGKEETSKEAKIRAKEK
                     RNRLSLLVQKPEFHEDTRSSRSGHLAKETRVSPEAVKWGESFDKLLSHRDGLEAFTR
                     FLKTEFSEENIEFWIACEDFKYKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEV
                     ITNSIQQTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRS
                     FTCNEFQDVQSDVAIWI"
     misc_feature    424..771
                     /note="RGS; Region: Regulator of G protein signalling
                     domain"
                     /db_xref="CDD:smart00315"
ORIGIN
```

```
Query Match   51.6%;  Score 601;  DB 9;  Length 2164;
Best Local Similarity   75.2%;  Pred. No. 1.4e-100;
Matches  816;  Conservative  0;  Mismatches  235;  Indels  34;  Gaps  4;

Qy     1  TTTTTGTAAGAAAAATCTGAGGAAAGAGATTCGGGATAGCGCTTTATTC-AGGATGTTTCC   59
          |||| ||||||||| |   ||| |   | ||  |||||    |||  ||| |||||||||||
Db    37  TTCTTTTTTGTAAACATTACTGTGTAAGAGTTGTGATAACTTTTTATTCTACTATGTATATG   96

Qy    60  TATGAAATAGCATTCATCGTGCGGAGAGAAGGACTAAGGAAATCTGACATCGTTGGT  119
          |||| |  |||||| |  |  | |||  | | | || | | ||||  |||||||
Db    97  TATGGAACAGTATTATTAATAAATGAACTAGGGAAGGGATGTAATAAATTAGACATCTTCCAT  156

Qy   120  CACTGGACACAGAATATGGATATGTCACTGGTTTTCTCTCCTCAATTAAATAATGTGTGAAT  179
```

```
VERSION     AX299975.1  GI:17129462
KEYWORDS
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Murray,D.L. and Gagnon,A.W.
  TITLE     Nucleic acids encoding a novel regulator of G protein signaling,
            rgs18, and uses thereof
  JOURNAL   Patent: WO 0183514-A 18 08-NOV-2001;
            Aventis Pharmaceuticals Products Inc. (US)
FEATURES             Location/Qualifiers
     source          1..1840
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match             51.5%;  Score 599.6;  DB 6;  Length 1840;
  Best Local Similarity   75.2%;  Pred. No. 2.5e-100;
  Matches  814;  Conservative    0;  Mismatches  234;  Indels   34;  Gaps    4;

Qy          4 TTGTAAGAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTC-AGGATGTTTCCTAT 62
Db         34 TTTTTTGTAAACATTACTGTAAGAGTTGTGATAACTTTTATTCTTACTACTATGTATATGTAT 93

Qy         63 GAAATAGCATTCATCTGTGGGAGGAGAAGACTAAGGAAATCTGACATCTGTTGGTCAC 122
Db         94 GGAATAGTATTATTAAATGAACTAGGGGAGGATGTAATAAATGTGACATCTTATCATT-- 151

Qy        123 TGGGACAGAATATGGATATGTCACTGGTTTCTTCTCTCAATTAAATATGTGAATCAA 182
Db        152 TTAGAGAAGAAGATGGAAACAACATTGCTTTCTTTTCTTTTCTCAATAAATATGTGAATCAA 211

Qy        183 AAGAGAAAACTTTTTCAAACTGGCGAAAGAAGAAGAACAAGCATCCGAGG 242
Db        212 AAGAAAAAACTTTTTCAAGTTAATACATGGTTAACATTGACATCCTGCACA 271

Qy        243 CCAAAATCAGACGAAGAAAAGAATAGACTAAGTCTTCTCCTACAGAGGCCTGACT 302
Db        272 CCAAAATCAGAGCTAAGGAAAAAAAAGAAATAGACTAAGTCTCTTGTGTCAGAAACCTGAGT 331

Qy        303 TCCATGGAGAGCCAGCTCAAGCCAGTGAATCTGCCCTCTGGCCAAAGAAACAGAGTCTCTC 362
Db        332 TCATGAAGACACACCCGCTCCAGTAGATCTGGGCACTTGGGCCAAAGAACAAGAGTCTCCC 391

Qy        363 CTTGAAGAAGCAGTGAAATGGCTGAAATCATTGACAAATTGCTCTCCTCCATAGAGATGGAG 422
Db        392 CTGAGAGAGCAGTGGGGAATGGGGGGTGAAACTGCTTTCCCATAGAATGGGAC 451

Qy        423 TGGATGTGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGGGG 482
Db        452 TAGAGGCTTTTACCAGATTTCTAAAACTGAATTCAGTGAAGAAGAAAATATTGAATTTTGGA 511

Qy        483 TCGCCTGAAGACTTCAAGAGAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAAGG 542
Db        512 TAGCCTGTGAAGATTTCAAGAGAATGCAAGGGACCTCAACAAATTCACCTTAAAGCAAAAG 571

Qy        543 CAATCTCATGAGAAATTCATTCAGAAATGATGCCCCCAAAGAGGTTAACATTGATTTTCATA 602
Db        572 CAATATATGAGAATTTATACAGACGTTATACAGAAGAGAGGTTAACCTGCATTTTCACA 631

Qy        603 CTAAAGAGAGTAATTGCTATAGCGAAATGCTAGAGACCATCGCCCAGCCACTCTCCACACGCAC 662
Db        632 CAAAAGAAGTCATTACAAACAGCATCACTCAACCTCACCTACCTCCACAGTTTGATGCGTGCAC 691

Qy        663 AAAGCAGAGTGACCAGCTACCAGTCATGGAACATGAGTCAAGGACCTGACAGTTATAAACGCTTTTTGAAATCTGAGA 722
Db        692 AAAGCAGAGTGATCAGCTGTCAGCAGAAGACAGTTATACACGTTTTCTGAAATCTGACA 751

Qy        723 CCTACTTACATTGATGAAGGAAGGAAGACCTCAGAGACGACCAACAAACCTTAGGAGGACGATCAC 782
```

```
Db        752 TCTATTTAGACTTGATGGAAGGAAGACCTCAGAGACCAACAAATCTTAGGAGGACGATCAC 811
Qy        783 GATCATTTACTTACACATGATTCCAAGATGTAAAGTCAGATGTTGCCATTGGTTATGAG 842
Db        812 GCTCATTTACCTGCAATGAATTCCAAGATGTACAATCAGATGTTGCCATTGGTTATAAA 871

Qy        843 TAAAAGTCATTTGTCTTCTTTGATGTGTATGTGTATATCTAAAATATATACTAATACT 902
Db        872 GAAAATTGATTTTGCTCATTTTATGACAAACTTATACATC-------- 912

Qy        903 AATGTGTACTTCTAAAATATAGCTTGTGTGTATAAGAAGAGATGATTTCATTTT-----TA 956
Db        913 ------TGCTTCTAACATATCGCATGTTTATGTTAAGAATTGGTCCATCCTTTAAACTG 966

Qy        957 AAATACCACCAGCAAATACATCATTAAAATAAATGTAAGAACTTTTTATATATACTAAAATAATT 1016
Db        967 AAATAATGTCATGTGAAATTATTTTAAAAATGTAAAACAAAACTTTCTGCTAACAAAATA 1026

Qy       1017 CATCATCTATCTTCCGAAATATTTTATGAAAAATCTATCTGATATTCTATTCTAATAAAAT 1076
Db       1027 CATACAGTATCTGCCAGTATATTCGTAAAAACCTTCTATTTGATGTCATTCCATTTATAA 1086

Qy       1077 TC 1078
Db       1087 TC 1088

RESULT 9
AX299976
LOCUS       AX299976                2144 bp    DNA     linear   PAT 26-NOV-2001
DEFINITION  Sequence 19 from Patent WO0183514.
ACCESSION   AX299976
VERSION     AX299976.1  GI:17129463
KEYWORDS
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Murray,D.L. and Gagnon,A.W.
  TITLE     Nucleic acids encoding a novel regulator of G protein signaling,
            rgs18, and uses thereof
  JOURNAL   Patent: WO 0183514-A 19 08-NOV-2001;
            Aventis Pharmaceuticals Products Inc. (US)
FEATURES             Location/Qualifiers
     source          1..2144
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match             51.5%;  Score 599.6;  DB 6;  Length 2144;
  Best Local Similarity   75.2%;  Pred. No. 2.5e-100;
  Matches  814;  Conservative    0;  Mismatches  234;  Indels   34;  Gaps    4;

Qy          4 TTGTAAGAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTC-AGGATGTTTCCTAT 62
Db         34 TTTTTTGTAAACATTACTGTAAGAGTTGTGATAACTTTTATTCTTACTACTATGTATATGTAT 93

Qy         63 GAAATAGCATTCATCTGTGGGAGGAGAAGACTAAGGAAATCTGACATCTGTTGGTCAC 122
Db         94 GGAATAGTATTAATGAACTAGGGGAGGATGTAATAAATGAACATCTCTTCATT-- 151

Qy        123 TGGGACAGAATATGGATATGTCACTGGTTTCTTCTCTCAATTAAATATGTGAATCAA 182
Db        152 TTAGAGAAGAAGATGGAAACAACATTGCTTTCTTTTCTTTTCTCAATAAATATGTGAATCAA 211

Qy        183 AAGAGAAAACTTTTTCAAACTGGCGAAAGAAGAAGAACAAGCATCCGAGG 242
Db        212 AAGAAAAAACTTTTTCAAGTTAATACATGGTTAACATTGACATCCTGCACA 271

Qy        243 CCAAAATCAGACGAAGAAAAGAATAGACTAAGTCTTCTCCTACAGAGGCCTGACT 302
```

```
TITLE       Cloning and characterization of a novel regulator of G protein
            signalling in human platelets
JOURNAL     Cell. Signal. 14 (7), 595-606 (2002)
MEDLINE     21952499
PUBMED      11955952
REFERENCE   2  (bases 1 to 2144)
AUTHORS     Gagnon,A.W., Murray,D.L. and Leadley,R.J. Jr.
TITLE       Direct Submission
JOURNAL     Submitted (16-MAY-2000) Pfizer Central Research, Eastern Point Rd,
            Groton, CT 06340, USA
FEATURES             Location/Qualifiers
     source          1..2144
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="platelet"
     CDS             163..870
                     /codon_start=1
                     /product="regulator of G-protein signaling 18"
                     /protein_id="AAK58589.1"
                     /db_xref="GI:14279411"
                     /translation="METTLLFFSQINMCESKEKTPFKLIHGSGKETSKEAKIRAKEK
                     RNRLSLLVQKPEFHEDTRSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLRAFTR
                     FLKIBFSEENIEFWIACEDFKKSKSGPQQIHLKARAIYEKFIQTDAPKEVNLDFHTKEV
                     ITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRS
                     FTCNEFQDVQSDVAIWL"
ORIGIN
Query Match           51.5%;  Score 599.6;  DB 9;  Length 2144;
Best Local Similarity 75.2%;  Pred. No. 2.5e-100;
Matches   814;  Conservative  0;  Mismatches  234;  Indels  34;  Gaps  4;
```

```
Qy     4 TTGTAAGAAAATCGAGGAAAGATTCGGGATAGCGGCTTTATTC-AGGATGTTTCCTAT   62
         ||||||        |   || ||||    |    |  ||||||  |||| || |||||
Db    34 TTTTTGTAAACATTACTGTAAGAGTTGTGATAACTTTTTATTTTATTCTACTATGTATATGTAT   93

Qy    63 GAAATAGCATTCATCGTGGGAGAGAGAGGACTAAGGAAATCTGACATCGTGGTCAC  122
         |       |||||||||           |||||     |   | |  |  ||  |
Db    94 GGAATAGTATTAATAAATGAACTAGGAGGATGTAATGTAATAAATTAGACAATCTCATT--  151

Qy   123 TGGGACAGAATATGGATATGTCACTGGTTTCTTCTTCAATTAAATATGTGAATCAA  182
         ||||     ||||||||| ||   ||||         ||||| ||||||||| |||
Db   152 TTAGAGAAGATGGAAACAACATGCTTTTCTTTTCTTTTCAAATAAATAATGTGAATCAA  211

Qy   183 AAGGAAAACTTTTTCAAATCAATGCGTCAGGGGTCAGGGAAAGAACAGCATCGAGG  242
         ||||||     |||  |||| ||     ||  ||||||     |||||||||||
Db   212 AAGAAAAAACTTTTTCAAGTTAATACATGTGGTTGCAGGAAAAAAGAAGAAGAACAAGCAAAGAAG  271

Qy   243 CCAAAATCAGAGCGAAAGAAAAAAGAATAGACTGATCGCCCTCTTGGCCAAAGAACCTGACT  302
         |||||||||||||     ||||||   |||||||||| ||||    ||| |||||||| ||
Db   272 CCAAAATCAGAGCTAAGGCAAAAAAGAAATAGACTAAGTCTTCTTGTCCAGAAACCTGAGT  331

Qy   303 TCCATGGAGAGACTCAAGCCAGTAGATCTGCCCTCTTGGCAAAGAACAAGAGTCTCT  362
         | ||| ||||||||||||||||||||||| |||| |||| |||||||||||||||||
Db   332 TTCATGAAGACACCCGCTCAGTAGATCTGGGCACTTGGCCAAAGAAACAAGAGTCTCCC  391

Qy   363 CTGAAGAAGCAGTGAAATGGGCTGAATCATTGACAAATTGCTTCTCTCATAGAGATGAG  422
         |||||||| |||||||||||| |||||||||||||| |||||  |  ||||||||| |
Db   392 CTGAAGAGGCAGTGAAATGGGTGAATCATTGACAAACTGCTTTCCCATAGAGATGGAC  451

Qy   423 TGGATGCTTTTACCAGAATTCTAAAACTGAATTCAGTGAGGAACATTGAATTTTGGG  482
         | |||||||||||| ||||||||||||||||||||||||||||  ||||||||||||
Db   452 TAGAGGCTTTTACCCAGAATTCTAAAACTGAATTCAGTGAGGAAGAAATATTGAATTTTGGA  511

Qy   483 TCGCCTGTGAAGACTTCAAGGACCTCAACAAATCATCCTAAAAGCAAAGG  542
         | |||||||||||| |||||    ||| |||||||||| ||||| |||||
Db   512 TAGCCTGTGAAGATTTCAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAAAGCAAAG  571

Qy   543 CAATCTATGAGAAATTCATTCAGAAATGCAGAGGTAACATTGATTTTCATA  602
         ||||| |||||||||| || ||||| ||||||  ||| |  ||||||| |
Db   572 CAATATATGAGAAATTTATACAGAACTGACTGAGTGCCCAAAAGAGGTTAACCTCGATTTTCACA  631

Qy   603 CTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGATACGGCAC  662
```

```
Db   663 AAAGCAGAGTGACCAGCTCATGGAACATGACAGTTTATAAACGCCTTTTGAAATCTGAGA  722
Db   692 AAAGCAGAGTGTATCAGCTCATGGAACAAGAGACTCAGAGCAACAAACCTTAGGACGATCAC  751
Qy   723 CCTACTTACATTTGATAGAAGGAAGAACCTCAGAGACCTCAGAGATGTAAGTCCAAGATGTTCCAAGATGTTATGAG  782
Db   752 TCTATTAGACTTGATGGAAGGAAGAACCTCAGAGACCTCAGAGATGTTATGAG  811
Qy   783 GATCATTACTTACTACAATGATTTCCAAGATGTAAAGTCAGATGTTGCCATTTGGTTATGAG  842
Db   812 GCTCATTTACCTGCAATGAATTCCAAGATGTACAATCAGATGTTGCCATTTGGTTTATAAA  871
Qy   843 TAAAAGTCATTGCTCTTGTCTTTTGATAGTGTGTATCTAAAATATATACTAATAACT  902
Db   872 GAAAATTGATTTTGCTCATTTTTATGACAAACTTATACATC---------  912
Qy   903 AATGTGTACTTCTAAAATATAGCTTGTGTATAAGAGAGATGATTTCATTTT------TA  956
Db   913 ------TGCTTCTAACATATCGCATTTTATGTTAAGATTTTGGTCCCATCCTTTAAACTG  966
Qy   957 AAATACACACATGCAAATATACATATATTAAGAACTTTTTTATATTATACTAAAAATAATT  1016
Db   967 AAAATGTCATGTGAAATTTATTTTTAAAAAATGAAATGTAAAAACAAAACTTTCTGCTAACAAAATA  1026
Qy  1017 CATCATCATCTTCCGAAATATATTTTATGGAAAATCTATCTGATATTCTATTCTAATAAAAT  1076
Db  1027 CATCAGTAGTATCTGCCAGTATATTCGTATAATTCTGTAAAACCTTCATTGATGTCATTCATTTATAA  1086
Qy  1077 TC  1078
Db  1087 TC  1088
```

```
RESULT 10
AF268036
LOCUS       AF268036                2144 bp    mRNA    linear   PRI 17-JUL-2002
DEFINITION  Homo sapiens regulator of G-protein signaling 18 mRNA, complete
            cds.
ACCESSION   AF268036
VERSION     AF268036.1  GI:14279410
KEYWORDS    .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2144)
AUTHORS     Gagnon,A.W., Murray,D.L. and Leadley,R.J. Jr.
```

```
                                          ITNSITQPTLHSFDAAQSRVYQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRSRS
                                          FTCNEFQDVQSDVAIWL"

ORIGIN

          Query Match       51.4%;   Score 597.8;   DB 9;   Length 2138;
          Best Local Similarity   74.6%;   Pred. No. 5.3e-100;
          Matches  806;   Conservative   0;   Mismatches 242;   Indels   33;   Gaps    3;

Qy        4  TTGTAAGAAAATCTGAGGAAAGATTCGGATAGGGCTTTATTCAGGATGTTTTCCTATG    63
                |||||||| ||| | |||||    |  |  ||  |||||||||||||    |||||||||||
Db       23  TTTTTGTAAACATTACTGTTATTATTGTGATAACTTTTATTCTACTATGTATATGTATG    82

Qy       64  AAATAGCATTCATCTGTGGAGAGAAGGACTAAGGAAATCTGACATCTGTTGGTTCACT   123
                ||||||||||||||    |||||||||||    |||||||||||    |||||||||||
Db       83  GAATAGTTAATAAAATGAACTAGGGAAGGATGTAATAAAATTAGACATCTTCAT--TT   140

Qy      124  GGGACAGAAATGGATATGTCACTCGGTTTCTTCTCCTCATTAAAATATGTGAATCAAA   183
                |||||||||||||||    |||||||||||    ||||||||||    ||||||||||
Db      141  TAGAGGAAGAGGGAAACAACATTGCTTTCTTTTCTTTTCTCAAATAAAATATGTGAATCAAA   200

Qy      184  AGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGGAAGAAGAAGAAGCATCGAGGC   243
                ||||||||||||||    |||||||||||    ||||||||||||    ||||||||||
Db      201  AGAAAAAACTTTTTCAAGTTAATACATGGTTCAGGGAAGAAAAGAAGAAGCAAAGAAGC   260

Qy      244  CAAAATCAGAGCGAAAGAAGAAAAAAAGGAATAGACTAAGTCTTCCTCCACAGAGGCCTGACTT   303
                ||||||||||||||    |||||||||||    ||||||||||||    |||||||||||
Db      261  CAAAATCAGAGCTAAGGAAGAAAAAAATGAAAAGAATAGACTAAGTCTTCTGTGCAGAAAACCTGAGTT   320

Qy      304  CCATGAGAGACTCAAGCACAGTAGATCTGCCCTTGGGCCAAAGAAACAAGAGTCTCTCC   363
                ||||||||||||||    |||||||||||    ||||||||||||    |||||||||||
Db      321  TCATGAGAACACCCGCTCCAGTAGATCTCGGGCACTCGGGCCAAAGAAACAAGAGTCTCCCC   380

Qy      364  TGAAGAGCAGTGAAATGGGCTGAATCATTGACAAATTGCTCTCCTCATAGAGATGGAGT   423
                ||||||||||||||    |||||||||||    ||||||||||||    |||||||||||
Db      381  TGAAGAGGCAGTGAAATGGGCTGAATCATTTGACAAACTGCTTTCCCATAGAGATGGACT   440

Qy      424  GGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTGGGT   483
                |||||||||||||    |||||||||||    |||||||||||    |||||||||||
Db      441  AGAGGGTTTTTACCAGATTTCTTAAAACTGAATTCAGTGAAGAAAATATTGAATTTTGGAT   500

Qy      484  CGCCTGTGAAGACTTCAAGAAATGCAAGAGGACCTCAACAAATCATCCTAAAAGCAAAGGC   543
                ||||||||||||||    |||||||||||    ||||||||||||    |||||||||||
Db      501  AGCCCTGTGAAGATTTCAAGAAATGCAAGGAGACCTCAACAAATTCACCTTAAAGCAAAAGC   560

Qy      544  AATCTATGAGAAATCATTCAGAATGATGCCCCCAAAGAGAGTTAACATTGATTTTCATAC   603
                ||||||||||||||    |||||||||||    |||||||||||    |||||||||||
Db      561  AATAATATGAGAAATTTATACAGAGTAATGCTGATGCCCCCAAAGAGAGTTAACCTTGATTTGATTTTCACAC   620

Qy      604  TAAAGAAGTAATTGCTAAGGACATCGCGCCCAGCCACTCTCCAGTTTTGATACGGCACA   663
                ||||||||||||||    |||||||||||    |||||||||||    |||||||||||
Db      621  AAAAGAGAGTCATTACAACAGCATCATCACCCTCCAACCTACCCTCCACAGTTTCATGTGTGCACA   680

Qy      664  AAGCAGAGGAGTGTACCGGCTGCCATGGAGCATGACCAGTTATAAGACGCTTTTCTGAGAC   723
                ||||||||||||||    |||||||||||    ||||||||||||    |||||||||||
Db      681  AAGCAGAGGAGTGTATCAGCTCAAGCATGCCATGGAACAAGACAGTTATACACGTTTTCTGAAATCTGACAT   740

Qy      724  CTACTTACATTGATGAGGAAGGACCTCAGAGACCAACAAAACTTAGGAGGACGATCACG   783
                ||||||||||||||    |||||||||||    |||||||||||    |||||||||||
Db      741  CTATTTAGACTTGATGAGGAAGGACCTCAGAGACCAACAAAATCCTAGGAGGACGATCACG   800

Qy      784  ATCATTTACTTACAATGATTCCAAGATGTAAAGTCAGATGTTGCCATTGGTTATGAGAGT   843
                ||||||||||||||    |||||||||||    ||||||||||    |||||||||||
Db      801  CTCATTTACCTGCAATGAATTCCAAGATGTTGCCATTCAGATGTTGCCATTGGTTATAAAAG   860

Qy      844  AAAAGTCATTTGCTCTCTTTGATATGTTTATGTGACAAACTTATACATC---------   903
                |||||||||||||    |||||||||||    |||||||||||    |||||||||||
Db      861  AAAATTGATTTTGCTCATTTTATGCACAAAACTTATACATC---------           900

Qy      904  ATGTGTACTTCTAAAATAGCTTGTGTATTAAGAAGAGATGAATTTCATTTT-----TAA   957
                ||||||||||||||    |||||||||||    ||||||||||    |||||||||||
Db      901  -----TGCTTCTAACATCGCATGTTTATGTTAAGATTTGGTCCCATCCTTTTAAACTGA   955

Qy      958  AATACACCATGCAAATACATATTAAATGTAAGAACTTTTTTATATTATACTAAAAATAAATTC   1017
```

Db   956 AATATGTCATGTGAAATTATTTTAAAAATGTAAAACAAACTTTCTGCTAACAAAATAC 1015
           |||| |||| || || ||| |
Qy   1018 ATCATCTATCTTCCGAAATATTTATGAAAATCTATCTGATATTCTATTCTAATAAAATT 1077
               | ||||| |||| | ||
Db   1016 ATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTATAAT 1075

Qy   1078 C 1078
           |
Db   1076 C 1076

RESULT 12
AR165090
LOCUS   AR165090   2217 bp   DNA   linear   PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6274362.
ACCESSION   AR165090
VERSION   AR165090.1 GI:16238486
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 2217)
  AUTHORS   Hodge,M.R. and Yowe,D.
  TITLE   RGS-containing molecules and uses thereof
  JOURNAL   Patent: US 6274362-A 1 14-AUG-2001;
FEATURES   Location/Qualifiers
     source   1..2217
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match   51.2%; Score 596; DB 6; Length 2217;
Best Local Similarity   75.8%; Pred. No. 1.1e-99;
Matches 791; Conservative 0; Mismatches 220; Indels 33; Gaps 3;

Qy   41 TTTATTCAGGAGGATGTTTTCCTATGAAATAGCATTCATCTGTGGGAGAGGAGGACTAAGG 100
           |||||||| ||| | | |||||||| |||| || |||||||||| |||||||
Db   69 TTATTCTACTATGTATATGTATGGAATAGTATTAATAAATGAACTAGGGAAGGATGTAAT 128

Qy   101 AAATCTGACATCTGTTGGTCACTGGGACAGAATATGGATATGTCACTGGTTTTCTTCTCT 160
           |||| || ||| |||||| |||| || |||| ||| ||||| | |||||| |
Db   129 AAATTAGACATTCTTCAT--TTTAGAGAGAAGATGAAACAACATTGCTTTTCTTTCTTTCT 186

Qy   161 CAATTAAAATATGTGAATCAAAAGAGAAAACTTTTTCAAACTAATACATGGGTCAGGG 220
           || | |||||||| |||||| ||| |||||||||||||| | ||| ||||||||
Db   187 CAAATAAAATATGTGAATCAAAAGAAAAAACTTTTTCAAGTTAATACATCATGGTTCAGGA 246

Qy   221 AAAGAGAAACAAGCATCGAGGCGCAAAATCAGAGCGAAAGAAAAAAGGAATAGGACTAAGT 280
           ||||| ||||||||||| || |||||| ||| | ||| || | |||| |||| |||
Db   247 AAAGAAGAAACAAGCAAAGAAGCAAAATCAGAGCTAAGGAAAAAAGAAATAGGACTAAGT 306

Qy   281 CTTCTCCTACAGAGGGCGCCGACTTCCATGGACAGCTCCAAGCCAGTAGATCTGCCCCTTG 340
           ||||| |||| || ||||| | | |||| |||| | |||||||| | |||||| |||
Db   307 CTTCTTGTGCAGAAAACCTGAGTTTCATGAAGACACCCGCTCCAGTAGATCTGGGCACTTG 366

Qy   341 GCCAAAGAAACAAGAGTCTCTCCTGAGAAGCAGCAGTGAAATGGGCTGAATCATTGACAAA 400
           |||||| ||||||||| | | |||||| |||||||| |||| |||| |||||||| ||
Db   367 GCCAAAGAAACAAGAGTCTCCCCTGAGAAGCAGCAGTGAAATGGGTGGTGAATCATTGACAAA 426

Qy   401 TTGCTCTCTCATAGAGATGGAGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 460
           ||||| | ||| || | |||| | | ||| |||| ||||| ||| |||| |||||||||
Db   427 CTGCTTCTTCCCATAGAGATGGACTAGAGGCTTTTTCACCAGATTTTTTAAAACTGAATTCAGT 486

Qy   461 GAGGAGAACATTGAATTTGGGTGCGCCGCTGAAGACTTCAAGAAATGCAAGAACCTCAA 520
           |||||| | |||||||||| | || | || ||| |||| || |||| |||||| |||
Db   487 GAAGAAAATAATTGAATTTTGGATAGCCTGTGAAGATTTCAAGAAAAAGCAAGGACCTCAA 546

Qy   521 CAAATCATCCTAAAAAGCAAAAGCAATCTATGAGAAATTCATTCAGAATGATGCCCCCAA 580
           ||| | | ||| ||||||||| |||| || | |||| || |||| | || |||||
Db   547 CAAATTCACCTTAAAGCAAAAGCAATAATATGAGAAAATTTATACAGACTGATGTCCCCAAAA 606

Qy   581 GAGGTTAACATTGATTTTCATCTAAAGAAGAAGTAATTGCTAAGACATCATCGCCCCAGCCACT 640

RESULT 14
AR214935
LOCUS   AR214935   Sequence 1 from patent   2217 bp   DNA   linear   PAT 25-SEP-2002
DEFINITION   AR214935 Sequence 1 from patent US 6410240.
ACCESSION   AR214935
VERSION   AR214935.1   GI:23312888
KEYWORDS   .
SOURCE   Unknown.
ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 2217)
AUTHORS   Hodge,M.R. and Yowe,D.
TITLE   RGS-containing molecules and uses thereof
JOURNAL   Patent: US 6410240-A 1 25-JUN-2002;
FEATURES   Location/Qualifiers
source   1..2217
                     /organism="unknown"
                     /mol_type="genomic DNA"

ORIGIN

Query Match   51.2%;   Score 596;   DB 6;   Length 2217;
Best Local Similarity   75.8%;   Pred. No. 1.1e-99;
Matches   791;   Conservative   0;   Mismatches 220;   Indels 33;   Gaps   3;

Qy   41 TTTATTCAGGATGTTTCCTATGAAATAGCATTCATCTGTGGGAGAGAAGGACTAAGG 100
           |||||| |||| || ||| ||||||| |||| | | | |||||||||||||||||
Db   69 TTATTCTACTATGTATATGGAATAGTATTAATAAATGAACTAGGGAAGGATGTAAT 128

Qy   101 AAATCTGACATCTGTTGGTCACTGGACAGAATATGGATATGTCACTGGTTTTCTTCTCT 160
           |||| ||||||| ||||||| |||||||||| | |||||||||||||| ||||
Db   129 AAATTAGACATCTCTTCAT--TTTAGAGAAGAAGATGGAAACAACATTGCTTTTCTTTTCT 186

Qy   161 CAATTAAATATGTGAATCAAAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGG 220
           |||||||||||||||||||||||||| ||||||||||||||| |||||||||||||
Db   187 CAAATAAATATGTGAATCAAAGAGAAAAAACTTTTTCAAGTAATACATGGTTCCAGGA 246

Qy   221 AAAGAGAAACAAGCATCGAGGCCAAAATCAGAGGCGAAAGAAAAAAGGAATAGACTAAGT 280
           ||||||||||||||||||||||||||||||||||||| | |||||||||||||||||||
Db   247 AAAGAGAAACAAGCAAAGAGCCAAAATCAGAGCTAAGGGAAAAATAGAACTAAGT 306

Qy   281 CTTCCTCCACAGAGGCCTGACTTCCATGGAGACTCAAGCAGTAGATCTGCCCTCTTG 340
           |||| ||||||| |||||||||||||||| |||| ||||||||| ||| ||||
Db   307 CTTCTGTGCAGAAACCTGAGTTTCATGAAGCACCCCGCTCCAGTAGATCTGGGCACTTG 366

Qy   341 GCCAAAGAAACAAGAGTCTCTCCTGAAGAAGCAGTGAAAATGGCTGAATCATTTGACAAA 400
           |||||||||||||||||||| |||||||||||||||| | |||||||||||||||||||
Db   367 GCCAAAGAAACAAGAGTCTCCCATAGAGAACAAGCAGTGAAAATGGGTGAAATCATTTGACAAA 426

Qy   401 TTGTCTCTCAGAGAGATGGAGTGGATGCTTTTACCAGATTCTTAAAACTGAATTCAGT 460
           |||| || |||| | ||| |||||||||||||||||||||||| |||||||||||||||
Db   427 CTGCTTTGTGCAGAAACCTGAGTTTCATGAAGCACCCCGCTCCAGTAGATCTGGGCACTTG 486

Qy   461 GAGGAGACATTGAATTTTGGGTCGCCTGTGAAGACTTCAAGGAAATCCAAGGAACCTCAA 520
           ||||||||||||||||| ||||| ||||||||||| ||| |||| |||| |||||||
Db   487 GAAGAAAATATTGAATTTTGGATATGCCGTGTGTCATCAGTGTATCAGTGTATCAGTGTAT 546

Qy   521 CAAATCATCCTAAAAGCAAGGCAATCTATGAGAAATCATTCAGATGATGCCCCCAAA 580
           ||||||||||| ||||||||||||||| ||| || |||||| |||||||||| |||
Db   547 CAAATCACCTTAAAGCAAGGCAATCTATTAGACTTAGAGATTTATACGACGTAAGTAGT 606

Qy   581 GAGGTTAACATTGATTTTCATACTGTCTAAGAAGTAATTGCTAAGACATCGCGCCCCACT 640
           ||||||||||||||| || | |||| |||||||||| || || | ||| ||||||||||
Db   607 GAGGTTAACCTTGATTTTCACACAGAAGAAGTCATTACCAGCTCAGTGAACACGTGATTAT 666

Qy   641 CTCCACAGTTTGATACGGCACAAAGCAGCAGAGTTCATGGAACATGACAGTTAT 700
           |||||||||||| | || |||||||||||||||| |||||||| |||||||||||
Db   667 CTCCACAGTTTTGATGCTGCACAAAGCAGCAGAGTATCACAGCTCATCTCTCTA 726

Qy   701 AAACGCTTTTGAAATCTGAGACCTACTTACATTTGATAGAAGGAAGACCTCAGAGACCA 760
           ||| || |||| |||| |||| ||||| |||| | |||||||||||||||||||||||
Db   727 ACACGTTTTCTGAAATCTGACATTCATTAGTACTTAGAGACTGTATGATGATGGAGACCA 786

Qy   761 AAACGCTTTTGAGAGGACGATCACCATCATTTACTTACCTGTCTTAAGATGTAAAGTCA 820
           ||| || ||||||||| || | ||| | |||||||| || ||||||||||| ||||
Db   787 ACAAATCTTAGGAGACGACGATCACCGTCATTTACTTGTCTTAAGATGTAAAGTCA 846

Qy   821 GATGTTGCCATTTGGTTATGAGTAAAAGTCATTTGTCTTCTTTTGATAGTGTATGTGTAT 880
           |||||||||||||||||||||||||||| ||||||| |||||||||||| |||| |
Db   847 GATGTTGCCATTTGGTTTATAAGAAAAATTTGATTTTGCTCATTTTATGACAAACTTATAC 906

Qy   881 ATCTAAAAATATATACTAATGTGTACTTCTAAAATAATAGTCTTGTGTATAAGAGAA 940
                                    TAAAATACACCAGTCAAATACATTAAATGTTAAGAACTT 994
Db   907 ATC--------   941

Qy   941 GATGATTTCATTTT------TAAAATACACCAGTCAAATACATTAAATGTTAAGAACTT 994
           ||| ||||||| | | || | ||||| || | ||||| ||||||||
Db   942 TTTGGTCCCATCCTTAAACTGAATATGTCATGTGAAATTATTTTTAAAAATGTAAAAAC 1001

Qy   995 TTTATATTACTAAAAATAATTCATCATCTATCCGAAATATTTTATGAAAATCTATC 1054
           ||||| | || | | ||| ||||| | | | | |||||||| | | |||
Db   1002 TTTATATTACTAAAAATAATTCATCATCTATCCGAAATATTTTATGAAAATCTATC 1054

Qy   1055 TGATAATTCATTCTAATAAAATTC 1078
           ||||||||||||| || |||||||
Db   1062 TTTGATGTCATTCCATTTATAATC 1085

```
Qy   761  ACAAACCTTAGGAGGACGATCACGATCATTACTTACACATGATTCCAAGATGTAAAGTCA  820
          ||||| ||||||||||||||||| |||||| ||| |||| |||||||||||||||| | ||
Db   787  ACAAATCTTAGGAGGACGATCACGCTCATTTACCTGCAATGAATTCCAAGATGTACAATCA  846

Qy   821  GATGTTGCCATTTGGTTATGAGTAAAAGTCATTTGTCTTCTTTGATAGTGTATGTGTAT   880
          |||||||||||||||||| |   |||| |||| ||  |  ||| ||
Db   847  GATGTTGCCATTTGGTTATAAAGAAAATGATTTTGCTCATTTTATGACAAACTTATAC    906

Qy   881  ATCTAAAAAATATATACTAATACTAATGTGTACTCTAAAATATAGCTTGTGTATAAGAAGA  940
          |||
Db   907  ATC-----------------------TGCTTCTAACATATCGCATGTTTATGTTAAGA    941

Qy   941  GATGATTCATTT-----TAAAATACACCATGCAAATACATATTAAATGTAAGAACTT     994
          |||||||||||      |||  ||  |||   |||| ||  |||||||||||  |||
Db   942  TTTGGTCCATCCTTTAAACTGAAATAATGTCATGTGAAATTATTTTTAAAAATGTAAAAAC 1001

Qy   995  TTTATATTATACTAAAATAATTCATCATCTATCTTCCGAAATATTTTATGAAAATCTATC  1054
          ||| | || |  ||||| |||| ||  | | |    | |||| ||  |||||| || |
Db  1002  AAAACTTTCTGCTAACAAAATACATACAGTATCATCTGCCCAGTATATTCTCTAAAACCTTCTA 1061

Qy  1055  TGATATTCTATTCTCAATAAAATTC  1078
          | |   || ||||  |||| ||||
Db  1062  TTTGATGTCATTCCATTTATAAATC  1085
```

RESULT 15

AC102163

```
LOCUS       AC102163              218670 bp    DNA     linear   ROD 06-NOV-2003
DEFINITION  Mus musculus chromosome 1, clone RP23-431N20, complete sequence.
ACCESSION   AC102163
VERSION     AC102163.13  GI:38194330
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 218670)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 1, clone RP23-431N20
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 218670)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrin,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,V., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Schupback.N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,D., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 218670)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,W., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,X., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Wu,X.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 218670)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,X., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-NOV-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Nov 6, 2003 this sequence version replaced gi:37777395.
            All repeats were identified using RepeatMasker:
            Smit,A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----------------------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            ----------------------- Project Information
            Center project name: L18154
            Center clone name: 431_N_20
            -----------------------
FEATURES             Location/Qualifiers
     source          1..218670
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="1"
                     /map="1"
                     /clone="RP23-431N20"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     misc_feature    1..2529I
                     /note="wgs end_extension clone_end:SP6"
     repeat_region   489..687
                     /rpt_family="B3"
     repeat_region   954..2106
                     /rpt_family="Lx5"
     repeat_region   2254..2346
                     /rpt_family="RSINE1"
```

```
repeat_region        2399...2622
                     /rpt_family="L1_MM"
repeat_region        4806...4975
                     /rpt_family="MTE"
repeat_region        5410...5442
                     /rpt_family="(CA)n"
repeat_region        6137...6197
                     /rpt_family="(TG)n"
repeat_region        6277...6445
                     /rpt_family="B3A"
repeat_region        6629...6782
                     /rpt_family="MLT1J"
repeat_region        complement(6866...6934)
                     /rpt_family="ID1_MM"
repeat_region        6962...7012
                     /rpt_family="AT_rich"
repeat_region        7643...7669
                     /rpt_family="AT_rich"
repeat_region        8296...8326
                     /rpt_family="AT_rich"
repeat_region        8698...8739
                     /rpt_family="(CA)n"
repeat_region        complement(8904...9079)
                     /rpt_family="ORR1A2"
repeat_region        9483...10099
                     /rpt_family="Lx2"
repeat_region        10100...10124
                     /rpt_family="(CAAAA)n"
repeat_region        10125...11467
                     /rpt_family="Lx2"
repeat_region        11468...11503
                     /rpt_family="(CAAA)n"
repeat_region        11508...11653
                     /rpt_family="B1_MM"
repeat_region        11654...12831
                     /rpt_family="Lx2"
repeat_region        12835...14147
                     /rpt_family="Lx2"
repeat_region        14192...14237
                     /rpt_family="(GA)n"
repeat_region        15650...15693
                     /rpt_family="(CA)n"
repeat_region        15790...15875
unsure               /note="single clone coverage"
repeat_region        16080...16274
unsure               /note="single clone coverage"
                     complement(16106...17211)
repeat_region        /rpt_family="L1_MM"
                     16154...16178
unsure               /note="<30 qual single clone coverage"
                     16161
unsure               /note="probably C"
                     16206...16245
unsure               /note="<30 qual single clone coverage"
repeat_region        17435...17476
                     /rpt_family="(TA)n"
repeat_region        17658...18245
                     /rpt_family="Lx6"
repeat_region        18537...18561
                     /rpt_family="AT_rich"
repeat_region        18987...19347
                     /rpt_family="Lx2"
repeat_region        19746...19996
                     /rpt_family="Lx6"
repeat_region        19997...20038
                     /rpt_family="(ATG)n"
repeat_region        20039...20599
                     /rpt_family="Lx6"
repeat_region        21204...21770
                     /rpt_family="Lx5"
repeat_region        21774...21804
                     /rpt_family="(TTAA)n"
repeat_region        21861...21890
```

```
repeat_region        /rpt_family="(TA)n"
                     complement(22318...22377)
repeat_region        /rpt_family="L1MCa"
                     complement(22393...22529)
repeat_region        /rpt_family="Lx2"
                     23925...24099
repeat_region        /rpt_family="(TGG)n"
                     24212...24238
repeat_region        /rpt_family="(CAAA)n"
```

```
Query Match          46.0%;    Score 535.4;   DB 10;   Length 218670;
Best Local Similarity  99.8%;    Pred. No. 7.3e-89;
Matches  536;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    582 AGGTTAACATTGATTTTCATACTAAAGAGTAATTGCTAAGAGCATCGCCCAGCCCACTC 641
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 144975 AGGTTAACATTGATTTTCATACTAAAGAGTAATTGCTAAGAGCATCGCCCAGCCCACTC 145034

Qy    642 TCCACAGTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATA 701
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145035 TCCACAGTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATA 145094

Qy    702 AACGCTTTTGAAATCTGAGACCTACTTACATTTGATAGAAGGAAGACCTCAGAGACCAA 761
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145095 AACGCTTTTGAAATCTGAGACCTACTTACATTTGATAGAAGGAAGACCTCAGAGACCAA 145154

Qy    762 CAAACCTTAGGAGACGATCACGATCATTTACTTACAATGATTTCCAAGATGTAAAGTCAG 821
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145155 CAAACCTTAGGAGACGATCACGATCATTTACTTACAATGATTTCCAAGATGTAAAGTCAG 145214

Qy    822 ATGTTGCCATTGGTTATGAGTAAAGTCATTTGTCTTCTTTGATAGTGTATGTGTATA 881
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145215 ATGTTGCCATTGGTTATGAGTAAAGTCATTTGTCTTCTTTGATAGTGTATGTGTATA 145274

Qy    882 TCTAAAATATATATACTAATACTAATGTGTACTTCTAAAATATAGTCTGTGTATAAGAAGAG 941
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145275 TCTAAAATATATATACTAATACTAATGTGTACTTCTAAAATATAGTCTGTGTATAAGAAGAG 145334

Qy    942 ATGATTTCATTTTAAAATACCACCATGCAAATACATATTAAATGTAAGAAACTTTTTATAT 1001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145335 ATGATTTCATTTTAAAATACCACCATGCAAATACATATTAAATGTAAGAAACTTTTTATAT 145394

Qy   1002 TATACTAAAATAATTCTTTATTTCACAATAACAGTCAGTAAGAAGAAGCTTTGAA 1061
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145395 TATACTAAAATAATTCTTTATTTCACAATAACAGTCAGTAAGAAGAAGCTTTGATATT 145454

Qy   1062 CTATTCTAATAAAATTCTTTATTTCACAATAACAGTCAGTCAGTAAGAAGAAGCTTTGTA 1118
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145455 CTATTCTAATAAAATTCTTTATTTCACAATAACAGTCAGTCAGTAAGAAGAAGCTTTGTA 145511
```

Search completed: August 20, 2004, 09:59:00
Job time : 4616.24 secs

OM nucleic - nucleic search, using sw model

```
                   GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.
```

Run on: August 20, 2004, 01:52:23 ; Search time 482.676 Seconds
                              (without alignments)
                              10244.763 Million cell updates/sec

Title: US-09-894-749-3
Perfect score: 1164
Sequence: 1 ttttgtaagaaaatcga.........gtactagtggatccgagctc 1164
Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_29Jan04:*
           1: geneseqn1980s:*
           2: geneseqn1990s:*
           3: geneseqn2000s:*
           4: geneseqn2001as:*
           5: geneseqn2001bs:*
           6: geneseqn2002s:*
           7: geneseqn2003as:*
           8: geneseqn2003bs:*
           9: geneseqn2003cs:*
          10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1164	100.0	1164	3	AAA52090	Aaa52090 Murine RG
2	1164	100.0	1164	6	AAD42498	Aad42498 Murine re
3	602.6	51.8	1884	6	ABQ99387	Abq99387 Human cod
4	599.6	51.5	1840	6	AAS18339	Aas18339 5'-RACE c
5	599.6	51.5	2144	6	AAS18340	Aas18340 Full leng
6	596	51.2	2217	3	AAA52089	Aaa52089 Human RGS
7	596	51.2	2217	6	AAD42497	Aad42497 Human reg
8	560.6	48.2	848	6	ABQ98617	Abq98617 Human ORF
9	192.2	16.5	1486	6	AAS18331	Aas18331 cDNA from
10	177.6	15.3	241	6	AAS18334	Aas18334 cDNA enco
11	156.2	13.4	681	9	ADD34052	Add34052 Mouse mit
12	152.4	13.1	1345	6	ABT10881	Abt10881 Human bre
13	152.4	13.1	1345	6	ABK83834	Abk83834 Human CDN
14	152.4	13.1	1345	7	ACC46750	Acc46750 Human COP
15	152.4	13.1	1345	7	ACA56680	Aca56680 Signallin
16	152.4	13.1	1345	9	ADD14613	Add14613 Human src
17	152.4	13.1	1364	3	AAF16132	Aaf16132 Human pro
18	152.4	13.1	1382	9	ADB47313	Adb47313 Human cDN
19	152.4	13.1	1629	7	ABT42293	Abt42293 Toxicity
20	150.2	12.9	408	7	ACC43243	Acc43243 Nucleotid
21	150.2	12.9	1795	4	AAH76414	Aah76414 RGS prote
22	143.6	12.3	540	6	ABA92505	Aba92505 Human reg
23	143.6	12.3	597	5	AAF57422	Aaf57422 Human RGS
24	143.6	12.3	939	6	ABZ11471	Abz11471 Human pol
25	142	12.2	543	6	ABL99975	Abl99975 Rat disea
26	142	12.2	745	6	AAV38084	Aav38084 Human reg
27	142	12.2	817	6	ABL99979	Abl99979 Rat disea
28	142	12.2	1691	3	AAZ36910	Aaz36910 cDNA enco
29	142	12.2	1923	6	AAX51745	Aax51745 DNA encod
30	142	12.2	1923	6	ABQ92597	Abq92597 Human sec
31	142	12.2	2075	7	ABX74396	Abx74396 Human cDN
32	139	11.9	800	4	ACA56493	Aca56493 Human sig
33	139	11.9	1238	7	AAL26766	Aal26766 Human bre
34	139	11.9	2934	6	ABK47336	Abk47336 cDNA enco
35	139	11.9	3165	9	ABX63113	Abx63113 Human CDN
36	139	11.9	3167	9	ADD12660	Add12660 Human cDN
37	135.8	11.7	3722	9	ADB58370	Adb58370 Toxicity-
38	134.6	11.6	750	7	ACC43242	Acc43242 Nucleotid
39	133.8	11.5	420	7	ABX37064	Abx37064 Bovine ES
40	132.6	11.4	2638	4	AAH02909	Aah02909 Human she
41	132.6	11.4	2638	6	ABL66645	Abl66645 Lung canc
42	132.6	11.4	2638	6	ABL62703	Abl62703 Colon ade
43	132.6	11.4	2638	7	ACA56492	Aca56492 Human sig
44	132.6	11.4	3327	7	ADA13378	Ada13378 Human int
45	129.8	11.2	1227	7	ABX74448	Abx74448 Human cDN

ALIGNMENTS

```
RESULT 1
AAA52090
ID   AAA52090 standard; cDNA; 1164 BP.
XX
AC   AAA52090;
XX
DT   04-DEC-2000  (first entry)
XX
DE   Murine RGS protein coding sequence.
XX
KW   RGS; regulators of G-protein signaling; GTPase activating protein; GAP;
KW   G-alpha protein; cell adhesion; chemotaxis; vulnerary; immunosuppressor;
KW   anti-rheumatic; anti-arthritic; anti-diabetic; anti-inflamatory;
KW   cytostatic; hepatotropic; anti-anaemic; modulator; gene therapy; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             134..841
FT                   /*tag= a
FT                   /product= "RGS_protein"
XX
PN   WO200046236-A2.
XX
PD   10-AUG-2000.
XX
PF   04-FEB-2000; 2000WO-US002977.
XX
PR   04-FEB-1999;  99US-00244314.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Hodge MR, Yowe D;
XX
DR   WPI; 2000-532893/48.
DR   P-PSDB; AAY97154.
XX
PT   Novel regulator of G-protein signaling nucleic acids and polypeptides,
PT   useful as diagnostic and investigative tools and to treat G-protein
PT   signaling disorders.
XX
PS   Claim 1; Page 102-104; 105pp; English.
XX
CC   The RGS (regulators of G-protein signaling) protein genes, clones
CC   AAH16395 and m1975, were identified in human and murine spleen cDNA
CC   libraries, respectively. Both proteins have unique N- and C-terminal
```

```
CC    sequences. The C-terminal location of the RGS domain is consistent with
CC    RGSs known to act as GTPase activating proteins (GAPs) for G-alpha
CC    proteins. G-alpha-i linked receptors support rapid adhesion and directed
CC    migration of leukocytes and other cell types. The novel RGS proteins may
CC    be used to modulate cell adhesion and chemotaxis, e.g. for aiding wound
CC    repair. The RGS proteins, related cDNAs and anti-RGS antibodies are
CC    useful for modulation, diagnosis and treatment of immune and respiratory
CC    disorders
XX
SQ    Sequence 1164 BP; 406 A; 202 C; 222 G; 334 T; 0 U; 0 Other;

Query Match            100.0%;   Score 1164;  DB 3;  Length 1164;
Best Local Similarity  100.0%;   Pred. No. 4.7e-246;
Matches 1164;   Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTCAGGATGGTTTTCCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTCAGGATGGTTTTCCT 60

Qy         61 ATGAAATAGCATTCATCTGTGGAGAGAAGAAGAAGACTAAGGAAATCTGACATCTGTTGGTC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAAATAGCATTCATCTGTGGAGAGAAGAAGAAGACTAAGGAAATCTGACATCTGTTGGTC 120

Qy        121 ACTGGGACAGAATATGGATATGTCACTGGTTTTCTTCTCAATTAAATATGTGTGAATC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTGGGACAGAATATGGATATGTCACTGGTTTTCTTCTCAATTAAATATGTGTGAATC 180

Qy        181 AAAAGAGAAAAACTTTTTCAAACTAATGCATGGTCAGGGAAAGAAGAAACAAGCATCGA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAAGAGAAAAACTTTTTCAAACTAATGCATGGTCAGGGAAAGAAGAAACAAGCATCGA 240

Qy        241 GGCCAAAATCAGAGCGAAAGAAAAAAAGGAATAAGACTAAGTCTTCTCCTACAGAGGCCTGA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCCAAAATCAGAGCGAAAGAAAAAAAGGAATAAGACTAAGTCTTCTCCTACAGAGGCCTGA 300

Qy        301 CTTCCATGAGAGACTCAAGCCAGTAGATCTCGCCCTGGCCAAAGAAACAAGAGTCTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTTCCATGAGAGACTCAAGCCAGTAGATCTCGCCCTGGCCAAAGAAACAAGAGTCTC 360

Qy        361 TCCTGAAGAAGCAGTGAAATGGGCTGAATCATTGACAAATTGCTCTCTCATAGAGATGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCTGAAGAAGCAGTGAAATGGGCTGAATCATTGACAAATTGCTCTCTCATAGAGATGG 420

Qy        421 AGTGGATGCTTTTACCAGATTTCTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGTGGATGCTTTTACCAGATTTCTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG 480

Qy        481 GGTCGCCTGTGAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTCGCCTGTGAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAA 540

Qy        541 GGCAATCTATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTTCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCAATCTATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTTCA 600

Qy        601 TACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGATACGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGATACGGC 660

Qy        661 ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAAATCTGA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAAATCTGA 720

Qy        721 GACCTACTTACATTGATAGAGGAAGACCTCAGAGACCAACAAACCTTAGGAGACGATC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACCTACTTACATTGATAGAGGAAGACCTCAGAGACCAACAAACCTTAGGAGACGATC 780

Qy        781 ACGATCATTTACTTACTTACAATGATTCCAAGATGTAAAGTCAGATGTTGCCATTGGTTATG 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACGATCATTTACTTACTTACAATGATTCCAAGATGTAAAGTCAGATGTTGCCATTGGTTATG 840

Qy        841 AGTAAAGTCATTGTCTTCTTTTGATAGTGTATGTGTATCTAAAATATATACTAATA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGTAAAGTCATTGTCTTCTTTTGATAGTGTATGTGTATCTAAAATATATACTAATA 900

Qy        901 CTAATGTGTACTTCTAAAATATAGCTTGTGTATAAGAGAGATGATTCATTTTTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTAATGTGTACTTCTAAAATATAGCTTGTGTATAAGAGAGATGATTCATTTTTAAAT 960

Qy        961 ACACCATGCAAATACATATATGTAAGGAACTTTTATATTATTACTAAAATAATTCATC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACACCATGCAAATACATATTAAATGTAAGAACTTTTATTATTACTAAAATAATTCATC 1020

Qy       1021 ATCTATCTTCCGAAATATTTTATGAAAATCATCTGATAATTCATTCTTCAATAAAATTCTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATCTATCTTCCGAAATATTTTATGAAAATCATCTGATAATTCATTCTTCAATAAAATTCTT 1080

Qy       1081 TATTTCTACAATAACAGTCAGTAAGAAGAAGCTTTGAAGCCGAATTCCAGCACACTGGCG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TATTTCTACAATAACAGTCAGTAAGAAGAAGCTTTGAAGCCGAATTCCAGCACACTGGCG 1140

Qy       1141 GCCGGTACTAGTGGATCCGAGCTC 1164
              ||||||||||||||||||||||||
Db       1141 GCCGGTACTAGTGGATCCGAGCTC 1164
```

RESULT 2
AAD42498
```
ID   AAD42498 standard; cDNA; 1164 BP.
XX
AC   AAD42498;
XX
DT   15-NOV-2002 (first entry)
XX
DE   Murine regulator of G-protein signalling (RGS) protein, m1975 cDNA.
XX
KW   Murine; screening; RGS; regulator of G-protein signalling; pancreatitis;
KW   inflammatory disorder; allergy; Grave's disease; arthritis; sinusitis;
KW   respiratory disorder; asthma; pneumonia; therapy; immune disorder;
KW   haematological disorder; haematopoiesis; platelet-associated disorder;
KW   thrombocytopaenia; migration; invasive disorder; leukaemia; anaemia;
KW   erythrocyte-associated disorder; hepatitis; m1975; gene; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             134..841
FT                   /*tag= a
FT                   /product= "Murine RGS protein"
XX
PN   US6411240-B1.
XX
PD   25-JUN-2002.
XX
PF   04-FEB-2000; 2000US-00498959.
XX
PR   04-FEB-1999; 99US-00244314.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Hodge MR, Yowe D;
XX
DR   WPI; 2002-588886/63.
DR   P-PSDB; AAE25830.
XX
PT   screening assays for identifying agent that binds to human or mouse RGS
PT   (regulators of G-protein signaling) protein or its variant, or RGS
PT   protein encoded by homologous DNA sequences, or to host cell expressing
PT   the proteins.
XX
PS   Claim 15; Col 55-58; 42pp; English.
XX
CC   The invention relates to screening assays for identifying an agent that
CC   binds to: human or mouse RGS (regulators of G-protein signalling) protein
CC   or its variant, or RGS protein encoded by homologous DNA sequences; or a
CC   host cell expressing the RGS protein or its variant, or RGS protein
```

```
CC   encoded by homologous DNA sequences. The invention is useful for
CC   identifying an agent that binds to human or mouse RGS protein or its
CC   variant, or RGS protein encoded by homologous DNA sequences, or a host
CC   cell expressing the RGS protein or its variant, or RGS protein encoded by
CC   homologous DNA sequences. The agents identified using the invention are
CC   useful for modulating the activity of RGS proteins and thus useful for
CC   treating immune and inflammatory disorders (Grave's disease, allergy,
CC   arthritis), respiratory disorders (asthma, pneumonia, sinusitis),
CC   haematological disorders (haematopoiesis, migration), platelet-associated
CC   disorders (thrombocytopaenia), invasive disorders (leukaemia),
CC   erythrocyte-associated disorders (anaemia), pancreatitis, hepatitis etc.
XX   The present sequence is murine RGS protein, m1975 cDNA
SQ   Sequence 1164 BP; 406 A; 202 C; 222 G; 334 T; 0 U; 0 Other;

Query Match            100.0%;  Score 1164;  DB 6;  Length 1164;
Best Local Similarity  100.0%;  Pred. No. 4.7e-246;
Matches 1164;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1  TTTTTGTAAGAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTCAGGATGTTTTCCT     60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  TTTTTGTAAGAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTCAGGATGTTTTCCT     60

Qy     61  ATGAAATAGCATTCATCTGTGGGAGAGGACTAAGAAATCTGACATCTGTTGGTC        120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  ATGAAATAGCATTCATCTGTGGGAGAGGACTAAGAAATCTGACATCTGTTGGTC        120

Qy    121  ACTGGGACAGAATATGGATATGTCACTGGTTTCTCTCTCAATTAAATATGTGTGAATC     180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  ACTGGGACAGAATATGGATATGTCACTGGTTTCTCTCTCAATTAAATATGTGTGAATC     180

Qy    181  AAAAGAGAAAACTTTTTCAAACTAATGCATGGTCAGGGAAAGAAGAAACAAGCATCGA     240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  AAAAGAGAAAACTTTTTCAAACTAATGCATGGTCAGGGAAAGAAGAAACAAGCATCGA     240

Qy    241  GGCCAAAATCAGAGCGAAAGAAAGAAAAAGGAATAGACTAAGTCTTCCTACAGAGGCCTGA  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GGCCAAAATCAGAGCGAAAGAAAGAAAAAGGAATAGACTAAGTCTTCCTACAGAGGCCTGA  300

Qy    301  CTTCCATGGAGAGACTCAAGCCAGTAGATCGCCCCTTGGCCAAAGAAACAAGAGTCTC     360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  CTTCCATGGAGAGACTCAAGCCAGTAGATCGCCCCTTGGCCAAAGAAACAAGAGTCTC     360

Qy    361  TCCTGAAGAAGCAGTGAAATGGGCTGAATCATTGACAAATTGCTCTCTCATAGAGATGG    420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TCCTGAAGAAGCAGTGAAATGGGCTGAATCATTGACAAATTGCTCTCTCATAGAGATGG    420

Qy    421  AGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGACAATTGAATTTG   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  AGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGACAATTGAATTTG   480

Qy    481  GGTGCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  GGTGCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAA  540

Qy    541  GGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTTCA   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTTCA   600

Qy    601  TACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGATACGGC   660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  TACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGATACGGC   660

Qy    661  ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTGATACAGTTATAAACGCTTTTG   720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTGATACAGTTATAAACGCTTTTG   720

Qy    721  GACCTACTTACATTGATGAGGAAGAACCTCAGAGACCTCAGAGAGACCAACAAACCTTAGGAGACGATC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  GACCTACTTACATTGATGAGGAAGAACCTCAGAGACCTCAGAGAGACCAACAAACCTTAGGAGACGATC  780

Qy    781  ACGATCATTTACTTACTACAATGATTTCCAAGATGTGTTGCCATTTGGTTATG         840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  ACGATCATTTACTTACTACAATGATTTCCAAGATGTGTTGCCATTTGGTTATG         840

Qy    841  AGTAAAAGTCATTTGTCTTCTTTTGATAGTGTATGTGTATCTAAAATATACTAATA      900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  AGTAAAAGTCATTTGTCTTCTTTTGATAGTGTATGTGTATCTAAAATATACTAATA      900

Qy    901  CTAATGTGTACTTCTAAAATATAGCTGTGTGTATAAGAAGAGATGATTCATTTTTAAAAT   960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  CTAATGTGTACTTCTAAAATATAGCTGTGTGTATAAGAAGAGATGATTCATTTTTAAAAT   960

Qy    961  ACACCATGCAAATACATATTAAAATGAAGACTTTTTATATTATACTAAAATAATTCATC  1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  ACACCATGCAAATACATATTAAAATGAAGACTTTTTATATTATACTAAAATAATTCATC  1020

Qy   1021  ATCTATCTTCCGAAATATTTTATGAAAATCATCATCTGATATTCTATTCTAATAAAATTCTT  1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  ATCTATCTTCCGAAATATTTTATGAAAATCATCATCTGATATTCTATTCTAATAAAATTCTT  1080

Qy   1081  TATTTCTACAATAACAGTCAGTAAGAAGAGCTTTGAAGCCGAATTCCAGCCACACTGGCG  1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  TATTTCTACAATAACAGTCAGTAAGAAGAGCTTTGAAGCCGAATTCCAGCCACACTGGCG  1140

Qy   1141  GCCGGTACTAGTGGATCCGAGCTC  1164
           ||||||||||||||||||||||||
Db   1141  GCCGGTACTAGTGGATCCGAGCTC  1164


RESULT 3
ABQ99387
ID   ABQ99387 standard; cDNA; 1884 BP.
XX
AC   ABQ99387;
XX
DT   25-FEB-2003   (first entry)
XX
DE   Human coding sequence SEQ ID 120.
XX
KW   Human; expressed sequence tag; EST; chromosome 16;
KW   haematopoietic disorder; central nervous system disease; viral infection;
KW   peripheral nervous system disease; non-healing wound; infectious disease;
KW   immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW   fungal infection; autoimmune disorder; coagulation disorder; nootropic;
KW   antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW   cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW   immunostimulant; cerebroprotective; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200259260-A2.
XX
PD   01-AUG-2002.
XX
PF   16-NOV-2001; 2001WO-US042950.
XX
PR   17-NOV-2000; 2000US-00714936.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI   Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR   WPI; 2002-590824/63.
DR   N-PSDB; ABP64801.
XX
PT   New isolated polynucleotide, useful in research, diagnostic or
PT   therapeutic methods, e.g. preventing or treating disorders involving
PT   aberrant protein expression or biological activity.
XX
PS   Claim 1; SEQ ID NO 120; 394pp; English.
XX
CC   The present invention relates to novel human coding sequences (ABQ99268-
CC   ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC   therapeutic, diagnostic and research methods. The polynucleotides may be
```

CC   used in the field of molecular biology as hybridisation probes, primers
CC   for PCR, for chromosome and gene mapping, for the recombinant production
CC   of protein, or in generation of anti-sense DNA or RNA. The
CC   polynucleotides are useful in diagnostics as expressed sequence tags
CC   (ESTs) for identifying expressed genes or for physical mapping of the
CC   human genome. The proteins may be used as molecular weight markers, or as
CC   nutritional sources or supplements. The proteins may be used to maintain
CC   and expand cell population in a totipotential or pluripotential state
CC   useful for re-engineering damaged or diseased tissues, transplantation,
CC   manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC   polynucleotides and proteins are useful for preventing, treating or
CC   ameliorating disorders involving aberrant protein expression or
CC   biological activity, e.g. haematopoietic disorders, central/peripheral
CC   nervous system diseases, mechanical and traumatic disorders, non-healing
CC   wounds, immune deficiencies and disorders, infectious diseases caused by
CC   viral, bacterial or fungal infection, autoimmune disorders, allergic
CC   reactions and conditions, coagulation disorders, or cancer. The
CC   polynucleotide sequences of the invention were assembled from ESTs
CC   isolated mainly by sequencing by hybridisation, and in some cases,
CC   sequences obtained from one or more public databases. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1884 BP; 681 A; 314 C; 328 G; 561 T; 0 U; 0 Other;

Query Match   51.8%; Score 602.6; DB 6; Length 1884;
Best Local Similarity   75.3%; Pred. No. 1.2e-122;
Matches 817; Conservative 0; Mismatches 234; Indels 34; Gaps 4;

```
Qy      1  TTTTTGTAAGAAAATCTGAGAAAGATTCGGATAGCGCTTAATTC-AGGATGTTTTCC     59
           | |||||||||||  | ||| |||| | || | |  |  |  | ||    | ||
Db     25  TTCTTTTTTGTAAACATTACTGTAAAGTTGTGATGAACTTTTATTCTACTATGTATATG     84

Qy     60  TATGAAATAGCATTCATCTGTGGGAGAAGAGGACTAAGAGAAATCTGACATCTGTTGGT    119
           |||| ||||| | ||||  ||||| | || ||||   |  |||    ||
Db     85  TATGGAATAGTTATTAAATGAACTAGGGAAGGATGTAATAAAATAAAATATTAAAAAT     144

Qy    120  CACTCGGACAGAATATGGATATGTCACTGGTTTCTTCTCTCCAATTAAAATATGTGAAT    179
           | | | |   | | ||||| ||| |  ||||||| || |||||||||| ||||||||||
Db    145  T--TTAGGAAGATGGAAACAACATTGCTTTCTTTCTTTTCCAATTAAATATGTGAAT     202

Qy    180  CAAAAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGGAAGAAGAGAAACAAGCATCG    239
           |||||| | ||||||||||||  |||| ||||  |||| ||| || |||||||||| |
Db    203  CAAAAGAAAAACTTTTTCAAGTAATACATGCTTCAGGAATGGAAGAGAAGAAACAAGCAAAG    262

Qy    240  AGGCCAAATACAGAGCAAAGAAAAAAGGAATAACGTCTCTCCTACAGAGCCTG         299
           |||||||| | ||| |||||| ||||| | ||| || |||||||||||| | |
Db    263  AAGCCAAATCACGAGCTAAGGAGAAATAGACAAAATAGAAACCTG                 322

Qy    300  ACTTCCATGGAGACTCAAGCCAGTAGATCTGCCCCTTCGGCCAAAGAAACAAGAGTCT    359
           | |||||||||||||||||||||| |||||||||||  |||||||||||||||||||
Db    323  AGTTTCATGGAGACCACCCCGCTCCAGTAGATCTCGGGCACTTGGCCAAAGAAACAAGAGTCT    382

Qy    360  CTCCTGAAGAAGCAGTGAAATGGGCTGAAATCATTGACAAATTGCTCTCCATAGAGATG    419
           | |||||||||||||||| | |||||||||||||||||||||||| |  | ||||||||
Db    383  CCCCTGAAGAAGCAGTGAAATGGGGTGAAATCATTGACAAATTGACAAACTGCTTCCATAGAGATG    442

Qy    420  GAGTGGGATGCTTTACCAGATTCTTAAAACTTCAGTGAGGAGAACATTGAATTTT     479
           ||| || | ||||||||||||||| |||| |||||||||||||||||||| ||||
Db    443  GACTAGAGGCTTTTACCAGATTCTTAAAACTTCAGTGAGGAGAAAAATATTGAATTTT    502

Qy    480  GGGTCGCCTGTGAAGACTCAAGAAATGCAAGGAAACCTCAACAAATCATCCTAAAAGCAA    539
           |||| || |||||||||||||||||  ||||| |||| |||||||  ||  ||||||||
Db    503  GGATAGCCTGTGAAGATTCAAGAAATTCAAGAAAAGCAAGGGACCTCAACAAATTCACCTTAAAGCAA    562

Qy    540  AGGAATCCATGAGAAATTCATTCAGAATGATGCCCCCAAGAGAGGTTAACATTGATTTTC    599
           |  ||  | |||||| | | || |||||  |||  | |||||||||||||||||||| |
Db    563  AAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAAGAGGTTAACCTTGATTTTC    622

Qy    600  ATACTAAGAGAGTAATTGCTAAGAGCATCGCCCCAGCCCACCTCCACAGTTTGATACGG    659
           | |  |||||||| | | |||||| |||||| ||| || |||||||||||||||| |||
Db    623  ACACAAAGAGAGTCATTACAACAGAGCTCATTACAAACTCACCTCCCACAGTTTGATGCTG    682
```

```
Qy    660  CACAAAGCAGAGTGACCGCTCATGGGACCAGTACAGCAGTTATAAAGCGTTTTGAAATCTG    719
           ||||||||||||||| | ||||||||||||||||||||||||||   ||||| |||||||
Db    683  CACAAAGCAGAGTGATGCAGCTCATGGGACCAGTACAGCAGTTATACACGTTTTCGAAATCTG    742

Qy    720  AGACCTACTTACATTGATAGAAGGAAGACCTCAGAGACCCAACAAACCTTAGGAGACGAT    779
           |  | |  | ||| ||||| |||||||||||||||||| ||||||| |||| |||||||||
Db    743  ACATCCATTTAGACTTGATGGAAGGAAGACCTCAGAGACCAACAAATCTTAGGAGACGAT    802

Qy    780  CACGATCATTACTTACTACAATGATTCCAAGATGGTAAAGTCAGATGTTGCCATTGGTTAT    839
           ||||  |||||||||| | |||||| ||||||  |   | |||||||||||||||||||||
Db    803  CACGCCATTACTTACCTGCAATGAATTCCAAGATACAATCAGATGTTGCCATTGGTTAT     862

Qy    840  GAGTAAAAGTCATTTGTCTTCTTTGATACGTGATGTGTATATCTAAAATATATACTAAT    899
           |  |||||| |||| |||    |       |||      | |  ||  ||||||
Db    863  AAAGAAAAATGATTTTGCTCATTTTATGACAAACTTATACATC------            906

Qy    900  ACTAATGTGACTTCTAAAATATAGCTTGTGTATAAGAAGAGATGATTTCATTTT----    954
                     | |||||||||| |||| | |  ||  |  | ||
Db    907  -------TGCTTCTAACATATCCGCATGTTTATGTTAAGAATTTGGTCCCATCCTTTAAA    957

Qy    955  -TAAAATACACATGCAAATACATATTAAAATGTAGGAACTTTTTATATATACTAAAATA   1013
            |||||||  |  | ||||| ||| ||  ||||  |||  |||||||||||||| |||| |
Db    958  CTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAATGAACAAAACTTTCTGCTAACAA   1017

Qy   1014  ATTCATCATCTATCTTCCGAAAATATTTATGAAAATCATCTGATATTCTATTCTAATAA   1073
           |  |||  | | | ||||| ||| ||||||    | |||||||| |||   | | ||||
Db   1018  ATACATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGCATTCATTTA   1077

Qy   1074  AATTC   1078
           | ||
Db   1078  TAATC   1082
```

RESULT 4
AAS18339
ID   AAS18339 standard; DNA; 1840 BP.
XX
AC   AAS18339;
XX
DT   12-MAR-2002 (first entry)
XX
DE   5'-RACE clone DNA used to generate full length human RGS18 cDNA.
XX
KW   Human; regulator of G protein signalling; RGS18; arterial thrombosis;
KW   platelet activation dysfunction; myocardial infarction; stroke;
KW   coronary artery disease; cerebrovascular disease; unstable angina;
KW   deep vein thrombosis; systemic thromboembolism; anti-coagulant;
KW   invasive cardiac procedure; 5'-RACE; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183514-A2.
XX
PD   08-NOV-2001.
XX
PF   26-APR-2001; 2001WO-US013540.
XX
PR   28-APR-2000; 2000US-0200786P.
PR   02-AUG-2000; 2000GB-00018833.
XX
PA   (AVET) AVENTIS PHARM PROD INC.
XX
PI   Murray D., Gagnon AW;
XX
DR   WPI; 2002-055453/07.
XX
PT   Isolated regulator of G protein signaling polypeptide, useful for
PT   prevention/treatment of platelet activation dysfunction, such as arterial
PT   thrombosis, myocardial infarction, coronary artery disease and stroke.
XX
PS   Claim 1; Page 116-117; 127pp; English.
XX

CC   The present invention relates to the isolation of polynucleotide
CC   sequences, that encode a novel regulator of G protein signalling (RGS)
CC   polypeptide, RGS18, from human platelets. The invention also provides
CC   nucleotide primers and probes specific for an RGS18 nucleic acid. The
CC   sequences of the invention are useful for the manufacture of a medicament
CC   for the prevention or treatment of a platelet activation dysfunction,
CC   such as arterial thrombosis, myocardial infarction, coronary artery
CC   disease, stroke, cerebrovascular disease, unstable angina, deep vein
CC   thrombosis, systemic thromboembolism, and in invasive cardiac procedures
CC   for anti-coagulant purposes. The sequences of the invention can also be
CC   used for the detection of RGS18 nucleic acids and RGS18 polypeptides. The
CC   present sequence represents 5'-RACE clone cDNA used to generate the full
CC   length human RGS18 CDNA
XX
SQ   Sequence 1840 BP; 668 A; 309 C; 328 G; 535 T; 0 U; 0 Other;

Query Match   51.5%;   Score 599.6; DB 6;   Length 1840;
Best Local Similarity   75.2%;   Pred. No. 5.2e-122;
Matches 814;   Conservative   0;   Mismatches 234;   Indels   34;   Gaps   4;

Qy   4 TTGTAAGAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTC-AGGATGTTTCCTAT   62
           ||| ||| || |||||| ||| |||||| ||||| || ||| ||| | ||||| ||||
Db   34 TTTTTGTAAACATTACTGTAAGAGTTGTGATAACTTTTATTCTACTATGTATATGTAT   93

Qy   63 GAAATAGCATTCATCTGTGGGAGAGAAGAGGACTAAGGAAATCTGACATCTGTTGGTCAC   122
           || |||| | || |||||||||||| |||||| ||||| ||||||| ||||||| ||
Db   94 GGAATAGTATTAATAAATGAACTAGGGAAGGATGTAATAAAATTAGACATCTTCATT-- 151

Qy   123 TGGGACAGAAATATGGAATATGTCACTGGTTTTCTTCTCAATTAAATATGTGGAATCAA   182
           ||| ||||| ||||| ||| ||||| ||||||| |||| |||| ||| ||||| |||||
Db   152 TTAGAGAAGATGGAAACAACATTGCTTTCTTTTCTCAATAAAATATGTGGAATCAA   211

Qy   183 AAGAGAAAAACTTTTTCAAACTAATGCATCAGGGAAAGAAGAAACAAGCATGCGAGG   242
           ||||||||||||||||| ||||| ||| || ||||| ||||| ||||||||||||
Db   212 AAGAGAAAAACTTTTTTCAAGTTAATACATGGTTCAGGGAAAGAAGAACAAGCAAAGAAG   271

Qy   243 CCAAAATCAGAGCGAAAGGAAAGAAAAAAAAGGAATAGACTAAGTCTTCCTCCTAGACTAAG   302
           |||||||||||||| |||| |||||| | ||| ||| || ||||| |||| |||
Db   272 CCAAAATCAGAGCTAAGGAAAAAGAAATAGACTAAGTCTTCTTGTGTCAGAAACCTGAGT   331

Qy   303 TCCATGAGAGACTCAAGCCAGTAGATCTGCCCCTGTGGCCAAAGAAACAAGAGTCTCTC   362
           ||||||||| ||||||||||||||| |||| ||||||||||||| |||||||||
Db   332 TTCATGAAGACACACCCGCTCCAGTAGATCTGGGCACTTGGCCAAAGAAACAAGAGTCCCC   391

Qy   363 CTGAAGAAGCAGTGAAATGGGCTGAATCATTGACAAATTGCTCTCCATAGAGATGGAG   422
           ||||||||||||| |||| |||||||||||| ||||| ||| |||||||||||| ||||
Db   392 CTGAAGAAGCAGTGCAGTGAAATGGGGTGAATCATTGACAAACTGCTTTCCCATAGAGATGGAC   451

Qy   423 TGGATGCTTTTACCAGAATTCTTAAAACTGAATTCAGTGAGGAACATTGAATTTTGGG   482
           ||||||| |||||||||||||||||||||||||| |||||| |||| |||||||
Db   452 TAGAGGCTTTTACCAGAATTCTTAAAACTGAATTCAGTGAAGAAATATTGAATTTTTTGGA   511

Qy   483 TCGCCTGTGAAGACTTCAAGAAATGCAAGGACAACCTCAACAAACATCCTAAAAGCAAAGG   542
           || |||||||||||| |||| || ||||| |||||||||||| ||||| |||||||||
Db   512 TAGCCTGTGAAGATTTCAAGAAATGCAAGGACCTCAACAAATTCACCTTAAAGCAAAG   571

Qy   543 CAATCTATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTTCATA   602
           ||||| || ||||| |||||||||| |||||||||||||||||||||||||| |||||
Db   572 CAATATATGAGAAATTTTATACGAACTGATGCCCCCAAAGAGGTTAACCTCGATTTTCACA   631

Qy   603 CTAAAGAGTAATTGCTAAGAGCCATCGCCCAGCCCCACTCTCCACAGTTTTGATACGGCAC   662
           | ||||||| ||| ||| |||| || |||||||||||| |||||||||||||||||||||
Db   632 CAAAAGAGTCATTACAAACAGCATCACTCAACCTCACCCTCCACAGTTTGATGCTCAC   691

Qy   663 AAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTGAAATCTTGAGA   722
           |||||||||||| ||||||||||||||| |||||||| ||| ||||| |||| ||||||
Db   692 AAAGCAGAGTGTATCAGCTCATGGAACAACAGTTATACACGTTTTCTGAAATCTGACA   751

Qy   723 CCTACTTACATTGATGAAGGAGGAGACCTCAGAGACCAACAAACCTTAGGAGGATCAC   782
           ||| ||||| |||||||| || ||||||| ||||| ||||| ||||| ||| |||||
Db   752 TCTATTTGACTTGATGAAGGAAGAAGACCTCAGAGACCAACAAATCTTAGGAGCGGATCAC   811

Qy   783 GATCATTACTTACAATGATTCCAAGATGTAAAGTCAGAGATGTTGCCATTGGTTTGATGAG   842
           || ||||| || ||||| |||| |||| |||| || |||||||||||| || ||| ||||
Db   812 GCTCATTACCTGCAATGAATTCCAAGAATTCCAAGATGTTGCCATTGGTTGCATTGGTTATAA   871

Qy   843 TAAAAGTCATTTGCTTCTTTGATAGTGTATTGATGTGTATGTGTATATCTAAAAATATTACTAACTACT   902
           ||||||||| ||| |||||||||| ||||||||||| |||||||| |||| |||
Db   872 GAAAATTGATTTTGCTCATTTTATGCAACACTTATACATC------------   912

Qy   903 AATGTGACTTCTAAAAATAGCTTCTGTGTATAGAAGAGATGATTTCATTTT-----TA   956
           ||||||||| |||||||||||| ||||||||||||||||||| ||||||||| ||
Db   913 -----TGCTTCTAACATATCGCATGTTTATGTTAGTGTTAAGATTTGGTCCATCCTTAAACTG   966

Qy   957 AAATACACCATGCAAATACATATTATTAAATGTAAGAACTTTTTTATATTATACTAAAATAATT   1016
           |||| ||| |||||||| ||||||| ||||| ||||||||| |||||| || |||||||||
Db   967 AAATATGTCATGTGAAATATTTTAAAAATGTAAAAACAAAACTTTCTGCTAACAAAATA   1026

Qy   1017 CATCATCTACCTTCCGAAATATTTTTATGAAAATCTATCTGATATTCTATTCTATTTATAAAAT   1076
           ||||| ||| ||| ||||||| || || | |||||| ||||||| ||| |||||||||||||
Db   1027 CATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTATAA   1086

Qy   1077 TC 1078
           || |
Db   1087 TC 1088

RESULT 5
AAS18340
ID   AAS18340 standard; cDNA; 2144 BP.
XX
AC   AAS18340;
XX
DT   12-MAR-2002 (first entry)
XX
DE   Full length cDNA encoding human RGS18.
XX
KW   Human; regulator of G protein signalling; RGS18; arterial thrombosis;
KW   platelet activation dysfunction; myocardial infarction; stroke;
KW   coronary artery disease; cerebrovascular disease; unstable angina;
KW   deep vein thrombosis; systemic thromboembolism; anti-coagulant;
KW   invasive cardiac procedure; ss.
XX
OS   Homo sapiens.
XX
FH   Location/Qualifiers
FT   Key
FT   CDS   143..870
FT   /*tag= a
FT   /product= "RGS18 polypeptide"
XX
PN   WO200183514-A2.
XX
PD   08-NOV-2001.
XX
PF   26-APR-2001; 2001WO-US013540.
XX
PR   28-APR-2000; 2000US-0200786P.
PR   02-AUG-2000; 2000GB-00018833.
XX
PA   (AVET) AVENTIS PHARM PROD INC.
XX
PI   Murray DL, Gagnon AW;
XX
DR   WPI; 2002-055453/07.
DR   P-PSDB; AAU10749.
XX
PT   Isolated regulator of G protein signalling polypeptide, useful for
PT   prevention/treatment of platelet activation dysfunction, such as arterial
PT   thrombosis, myocardial infarction, coronary artery disease and stroke.
XX
PS   Claim 1; Fig 1; Fig 1; 127pp; English.
XX
CC   The present invention relates to the isolation of polynucleotide
CC   sequences, that encode a novel regulator of G protein signalling (RGS)
CC   polypeptide, RGS18, from human platelets. The invention also provides
CC   nucleotide primers and probes specific for an RGS18 nucleic acid. The

```
CC   sequences of the invention are useful for the manufacture of a medicament
CC   for the prevention or treatment of a platelet activation dysfunction,
CC   such as arterial thrombosis, myocardial infarction, coronary artery
CC   disease, stroke, cerebrovascular disease, unstable angina, deep vein
CC   thrombosis, systemic thromboembolism, and in invasive cardiac procedures
CC   for anti-coagulant purposes. The sequences of the invention can also be
CC   used for the detection of RGS18 nucleic acids and RGS18 polypeptides. The
CC   present sequence represents the full length cDNA sequence encoding RGS18
CC   from human platelets
XX
SQ   Sequence 2144 BP; 784 A; 340 C; 365 G; 655 T; 0 U; 0 Other;

Query Match             51.5%;  Score 599.6;  DB 6;  Length 2144;
Best Local Similarity   75.2%;  Pred. No. 5.4e-122;
Matches  814;  Conservative    0;  Mismatches  234;  Indels   34;  Gaps    4;

Qy          4 TTGTAAGAAAATCTGAGGAAGAGATTCGGGATAAGCGCTTATTC-AGGATGTTTCCTAT   62
              ||||||||| ||  ||||||   ||   ||||  || |||||||| ||||||||| ||||
Db         34 TTTTTTGTAAACATTACGTAAGAGTTGTGTAACATTTTTATTCTACTACGTATATGTAT   93

Qy         63 GAAATAGCATTCATCTGTGGGAGAGAAGAAGAAATCTGACATCTGTTGGTCAC        122
              |||||||||||  || ||||||||||||| ||  ||||||| | ||||||||||
Db         94 GGAATAGTATTATAATGAACTAGGGAAGGATGTAATAAATTAGACATCTCTTCATT--   151

Qy        123 TGGGACAGAATATGGATATGTCACTGGTTTTCTTCTCAATTAAATGTGTGAATCAA     182
              || | ||  ||| | |||||   ||||||  || |  ||||||||||||||||||||
Db        152 TTAGAGAAGAGATGGAAACAACATTGCTTTCTTTCTTTTCAATTAAATGTGTGAATCAA   211

Qy        183 AAGAGAAAAACTTTTTTCAAACTAATGCATGGGTCAGGGAAGAAGAAACAAGCATCGAGG  242
              ||||||||||||||  ||||| ||| | ||| |||| |||||||||||||||| |||||
Db        212 AAGAAAAAACTTTTTCAAGTTAACATGGTTCAGGAAAGAAAGAAGAAACAAGCAAAGAAG  271

Qy        243 CCAAAAATCAGAGCGAAAGAAAAAAGGAATAGACTAAGTACATTCTTCTCCCACAGGCGACT  302
              || ||||| |||||||||  || |  ||||||||| ||||| || | || ||||| ||||||
Db        272 CCAAAATCAGAGCTAAGGAAATAGACAAGAAGAAATAGACTAAGTCTCTTGTGCAGAAACCTGAGT  331

Qy        303 TCCATGGAGAGACTCAAGCCAGTAGATCTGCCCCTCTTGGCCAAAGAAACAAGAGTCTCTC  362
              ||||||||||||||  |||| |||||||| | ||||||||||||||||||||||||||||
Db        332 TTCATGAAGAACACCCGCTCCAGTAGATCTCGGGCACTTGGCCAAAGAAACAAGAGTCTCCC  391

Qy        363 CTGAAGAAGCAGTGAAATGGGCTGAATCATTGACAAATTGCTCTCTCATAGAGATGGAG     422
              |||||||||||||||||||||||||| ||||||| | | |||||||||| || ||||||
Db        392 CTGAAGAAGGCAGTGAAATGGGTGAATCATTTGACAAACTGCTTTCCCATAGAGATGGAC    451

Qy        423 TGGATGCTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGGG         482
              ||||||| || |||||||||| ||||||||||||| |||||||||||||||||
Db        452 TAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAAATATTGAATTTTTGGA   511

Qy        483 TCGCCCTGTGAAGAACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAAGG  542
              ||||||||||||||||||||||||||||||||||||| ||||| |||| |  || |||||||
Db        512 TAGCCTGTGAAGATTTCAAGAAATGCAAGGAACCTCAACAAATTCACCTTAAAGCAAAG     571

Qy        543 CAATCTCATGAGAAATTCATTCAGAAATGATGCCCCCAAAGAGTTAACATTGATTTTCATA    602
              ||| | |||||||||| || ||||||||| | ||||||||||||  ||||||||||  ||||
Db        572 CAATATATGAGAAAATTTATACAGACTGATGCCCCCAAAGAGTTAACCTGATTTTTCACA    631

Qy        603 CTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGATACCGCAC    662
              |||||||||||| | ||||||||| ||||||||| || || |||||||||||| |||| ||
Db        632 CAAAAGAAGTATCATTACAAACACAGCATCACAGAGCCTCCACAGTTTTGATGCGTGCAC   691

Qy        663 AAAGCAGAGTGTACCAGCTCATGGAACATGAACAGAACAAGACAGTTATACACGTTTCTGA  722
              ||||||||||||||||| ||||||||||||||| | |||| |||||| |  |||||||| |
Db        692 AAAGCAGAGTGTATCAGCTCATGGAACTAGGAACAGACAGTTATACACGTTTTCTGAAATCTGACA  751

Qy        723 CCTACTTACATTGATAGAAGGAGACCTCAGAGACCTTAGGAGACGACGATCAC   782
              |  |||| |  ||||| |||||||| |||||| |||||| | || | ||||||
Db        752 TCTATTTGACTTGATGGAAGGAAGACCTCAGAGACCTTAGGAGACGACGATCAC   811

Qy        783 GATCATTTACTTACAATGATTCCAAGATGTAAAGTCAGATGTGCCATTTGGTTATGAG    842
              |||||||| ||||| |||  |||||||| |||||||||||||| |||||||||||||||
Db        812 GCTCATTTACCTGCAATGAATTCCAAGATTACAAATCAGATGTTGCCATTTGGTTATAAA   871

Qy        843 TAAAAGTCATTTGTCTCTTTTTGATAGTGTATGTGTATGTGTATGTATATCTAAAATATATACTAATACT  902
```

```
Db        872 GAAAATTGATTTGCTCATTTTATGACAAACTTATACATC-----------------   912
              ||||| |||||  ||   |||||||     |||||||||
Qy        903 AATGTGTACTTCTAAAAATAGCTGTGTATAAGGAAGAGAGATGAGGATGATTCATTT-----TA   956
              |||||        ||||||| || |||||||| || | ||| || ||| |||||||     ||
Db        913 -----TGCTTCTACACATACGCATGTTATGTTAAGATTGGGTCCCATCCTTTAAACTG   966

Qy        957 AAATACACCATGCAAATACATATTCAAGTGAGATGTTTTTATATTATACTAAGATAATT  1016
              ||||| ||| |  || |||||  ||| || || |||||||||||| || | || | |||
Db        967 AAATATGTCATGTGAATTATTTTAAAAATGTAAAATGTAAAAACAAACTTCTGCTAACAAAATA  1026

Qy       1017 CATCATCTATCTTCCGAAAATATTTTATGAAAAATCTATCTGATATTCTATTCTAATAAAAT  1076
              |||||| || ||   |||||| |||||||   |||||| |||||||||  || ||||||||
Db       1027 CATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTGTCATTCCATTTATAA  1086

Qy       1077 TC  1078
              ||
Db       1087 TC  1088
```

```
RESULT 6
AAA52089
ID   AAA52089 standard; cDNA; 2217 BP.
XX
AC   AAA52089;
XX
DT   04-DEC-2000 (first entry)
XX
DE   Human RGS protein coding sequence.
XX
KW   RGS; regulators of G-protein signaling; GTPase activating protein; GAP;
KW   G-alpha protein; cell adhesion; chemotaxis; vulnerary; immunosuppressor;
KW   anti-rheumatic; anti-arthritic; anti-diabetic; anti-inflamatory;
KW   cytostatic; hepatotropic; anti-anaemic; modulator; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             160..867
FT                   /tag= a
FT                   /product= "RGS_protein"
XX
PN   WO200046236-A2.
XX
PD   10-AUG-2000.
XX
PF   04-FEB-2000; 2000WO-US002977.
XX
PR   04-FEB-1999;  99US-00244314.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Hodge MR,  Yowe D;
XX
DR   WPI; 2000-532893/48.
DR   P-PSDB; AAY97153.
XX
PT   Novel regulator of G-protein signaling nucleic acids and polypeptides,
PT   useful as diagnostic and investigative tools and to treat G-protein
PT   signaling disorders.
XX
PS   Claim 1; Page 100-102; 105pp; English.
XX
CC   The RGS (regulators of G-protein signaling) protein genes, clones
CC   AAH16395 and m1975, were identified in human and murine spleen cDNA
CC   libraries, respectively. Both proteins have unique N- and C-terminal
CC   sequences. The C-terminal location of the RGS domain is consistent with
CC   RGSs known to act as GTPase activating proteins (GAPs) for G-alpha
CC   proteins. G-alpha-i linked receptors support rapid adhesion and directed
CC   migration of leukocytes and other cell types. The novel RGS proteins may
CC   be used to modulate cell adhesion and chemotaxis, e.g. for aiding wound
CC   repair. The RGS proteins, related cDNAs and anti-RGS antibodies are
CC   useful for modulation, diagnosis and treatment of immune and respiratory
```

```
CC  disorders
XX
SQ  Sequence 2217 BP; 801 A; 359 C; 380 G; 677 T; 0 U; 0 Other;

Query Match     51.2%;   Score 596;  DB 3;  Length 2217;
Best Local Similarity  75.8%;  Pred. No. 3.4e-121;
Matches  791;  Conservative  0;  Mismatches  220;  Indels  33;  Gaps  3;

Qy       41 TTTATTCAGGATGTTTTCCATGAAATAGACATTCATCTGTGGGAGAAGAGGACTAAGG 100
            ||||||||||| ||||||||| |||||||||||||||||||| ||||||||||||||
Db       69 TTTATTCTACTATGTATGTATGGAATAGTATTATTAATTAAGAACTAGGGAAGGATGTAAT 128

Qy      101 AAAATCTGACATCGTTGGTCACTGGGACCAGAATATGGATATGTCA------------- 160
            ||||| ||||||||||||||||||| ||||||||| |||||||||
Db      129 AAATTAGACATCTTCAT--TTTAGAGAGAAGAGATGGAAACAACATTGCTTTCTTTTTCT 186

Qy      161 CAATTAAATATGTGAATCAAAGAGAGAAAACTTTTTTCAAACTAATACCATGGGTCAGGG 220
            ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      187 CAAATAAATATGTGAATCAAAGAGAGAAAAACTTTTTTCAGTTAATACACATGGGTTCAGGA 246

Qy      221 AAAGAAGAAACAAGCATCGAGCCAAAATCAGAGGCGAAAGAAAAAGGAATAGAACTAAGT 280
            ||||||||||| ||||||||||||||||||| |||||||||||||||||||||||||
Db      247 AAAGAAGAAACAAGCAAGAGCCAAGAAATCAGAGCTAAGAGAAAAAAAGAAATAGAACTAAGT 306

Qy      281 CTTCCTCCTACAGAGGGCCTGACTTCCATGGAGCGAGACTCAAGCCAGTAGATCTGCCCTCTTG 340
            ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      307 CTTCTTGTGCAGAAAACCTGAGTTTCATGAAGACACCCCGCTCCAGTAGATCTGGGCACTTG 366

Qy      341 GCCAAAGAAGAAAGAGTCTCTCCTGAAGAAGCCAGTGAAATGGGCTGAATCATTTGACAAA 400
            ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db      367 GCCAAAGAAACAAGAGTCTCCCCTGAAGAGGCGAGTGAATCATTTGAATCATTTGACAAA 426

Qy      401 TTGCTCTCTCATAGAGATGGAGTGGATGCGTCTTTTACCAGATTTCTTAAAACTGAATTCAGT 460
            |||||||||||||||||||||||||||| ||||||| |||||||||||||||||||||||
Db      427 CTGCTTTTCCCATAGAGATGGAGTGGACTAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 486

Qy      461 GAGGAGAACATTGAATTTGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAA 520
            |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db      487 GAAGAAAAATATTGAATTTGGATAGCCTGTGAAGACTTCAAGAAAAGCAAGGGACCTCAA 546

Qy      521 CAAATCATCCTAAAAGCAAAGCAACTCATGAGAAATTCATTCAGAATGATGCCCCCAAA 580
            ||||||||||||||||||||||||| |||||||||| ||||||||||||||||||||||
Db      547 CAAATTCACCTTAAAGCAAAGCAAAAGCAATATATGGAAAATTTATACAGACTGATGCCCCAAAA 606

Qy      581 GAGGTTAACATTTGATTTTCATACTAAGAAGGAAATTGCTAAGAAGTAAAATGCGCCAGCCCACT 640
            ||||||||||||||||||||||| |||||||||||||||||||||| ||||||||||||
Db      607 GAGGTTAACCTTTGATTTTCATTACAAAAGAAGTCATTACAAACAGCATCAACCTTACC 666

Qy      641 CTCCACAGTTTTGATACGGCCACAAAGCAGAGTGTACCAGCTCATGGAACATGCAGTTAT 700
            ||||||||||||||||||||||||||||||||| ||||||||||||| |||||||||||
Db      667 CTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGACAAGACAGTTAT 726

Qy      701 AAACGCTTTTGAAAATCTGAGACTACTTACATTTGATAGAAGGAAGACCTCAGAGACCA 760
            ||||| |||||||||||||||||||||| ||||||||||||||| |||||||||||||||
Db      727 ACACGTTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAAGACCTCAGAGACCA 786

Qy      761 ACAAACCTAGGAGACGATCACGATCATCATTACTTACTACAGATTTCCAAGATGTAAAGTCA 820
            |||||| |||||||| |||||||||||||||| |||||||| |||||||||||||||||
Db      787 ACAAATCTAGGAGACGATCACGCTCATCATCCTCATTTACCTGCAATGAATTCCAAGATGTACAATCA 846

Qy      821 GATGTTGCCATTGGTTATGAGTAAAAGTCATTTGTCTTCTTTTGATAGTGTATGTGTAT 880
            |||||||| |||||||||| |||||||||||| |||||||| ||||||| ||||||||
Db      847 GATGTTGCCATTTGGTTATAAAGAAAATTGATTTTGCTCATTTTTATGACAAACTTATAC 906

Qy      881 ATCTAAAATATATACTAATAACTAATGCTGTACTTCTAAAATAATAGCTTGTGTATAAGAGAGA 940
            ||||||||||||||||| |||||||||||||||||| ||||||||||||||||||||||
Db      907 ATC-----------------------TGCTTCTAACATATCGCATGTTTATGTTAAGA 941

Qy      941 GATGATTTCATTTT------TAAAATACACCATGCAAATACATATTAAATGTAAGAACTT 994
            ||||||||||||||      ||||||||||||||||||||||||||||||||||||||||
Db      942 TTTGGTCCCATCCTTTAACCTGAATATGAATATGCATGTGAAATATTTTTAAAAATGTAAAAAC 1001

Qy      995 TTTATATTATACTAAAATAATTCATCATCTATCTTCCGAAATATTTTATGAAAAAATCTATC 1054
            || ||| || ||| | ||| ||||||| || ||||| |||||||| | || || |||||||
Db     1002 AAAACTTTCTGCTACACAAAATACACAGTACACATACACAATTC 1078

Qy     1055 TGATATTCTATTCTAATAAAAATTC 1078
            | ||||||| |||| || ||| |||||
Db     1062 TTTGATGTCATTCCATTTATAATC 1085
```

```
RESULT 7
AAD42497
ID   AAD42497 standard; cDNA; 2217 BP.
XX
AC   AAD42497;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Human regulator of G-protein signalling (RGS) protein, h16395 cDNA.
XX
KW   Human; screening; RGS; regulator of G-protein signalling; pancreatitis;
KW   inflammatory disorder; allergy; Grave's disease; arthritis; sinusitis;
KW   respiratory disorder; asthma; pneumonia; therapy; immune disorder;
KW   haematological disorder; haematopoiesis; platelet-associated disorder;
KW   thrombocytopaenia; migration; invasive disorder; leukaemia; anaemia;
KW   erythrocyte-associated disorder; hepatitis; h16395; gene; ss.
XX
OS   Homo sapiens.
XX
FH                     Location/Qualifiers
FT   CDS               160..867
FT                     /*tag=  a
FT                     /product= "Human RGS protein"
XX
PN   US6410240-B1.
XX
PD   25-JUN-2002.
XX
PF   04-FEB-2000;  2000US-00498959.
XX
PR   04-FEB-1999;  99US-0244314.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Hodge MR, Yowe D;
XX
DR   WPI; 2002-588886/63..
DR   P-PSDB; AAR25829.
XX
PT   screening assays for identifying agent that binds to human or mouse RGS
PT   (regulators of G-protein signalling) protein or its variant, or RGS
PT   protein encoded by homologous DNA sequences, or to host cell expressing
PT   the proteins.
XX
PS   Claim 15; Col 51-56; 42pp; English.
XX
CC   The invention relates to screening assays for identifying an agent that
CC   binds to: human or mouse RGS (regulators of G-protein signalling) protein
CC   or its variant, or RGS protein encoded by homologous DNA sequences; or a
CC   host cell expressing the RGS protein or its variant, or RGS protein
CC   encoded by homologous DNA sequences. The invention is useful for
CC   identifying an agent that binds to human or mouse RGS protein or its
CC   variant, or RGS protein encoded by homologous DNA sequences, or a host
CC   cell expressing the RGS protein or its variant, or RGS protein encoded by
CC   homologous DNA sequences. The agents identified using the invention are
CC   useful for modulating the activity of RGS proteins and thus useful for
CC   treating immune and inflammatory disorders (Grave's disease, allergy,
CC   arthritis), respiratory disorders (asthma, pneumonia, sinusitis),
CC   haematological disorders (haematopoiesis, migration), platelet-associated
CC   disorders (thrombocytopaenia), invasive disorders (leukaemia), anaemia),
CC   erythrocyte-associated disorders (anaemia), pancreatitis, hepatitis etc.
CC   The present sequence is human RGS protein, h16395 cDNA
XX
SQ   Sequence 2217 BP; 801 A; 359 C; 380 G; 677 T; 0 U; 0 Other;
```

```
Query Match    51.2%;    Score 596;  DB 6;  Length 2217;
Best Local Similarity   75.8%;  Pred. No. 3.4e-121;
Matches   791;  Conservative   0;  Mismatches 220;  Indels  33;  Gaps   3;

Qy    41 TTTATTCAGGATGTTTTCCTATGAAATAGCATTCATCGTGTGGGAGAGAAGGACTAAGG 100
          |||||||  || |  ||  || || ||||  ||| |  |  ||  |     | ||| ||
Db    69 TTTATTCTACTACTATATGTATGTGGAATAGTATTAATAAATGAACTAGGGATGTAAT 128

Qy   101 AAAATCTGACATCTGTGGTCACTGGGACAGAATATGATATGTCACTGGTTTCTTCTCT 160
          ||| | |||||||| ||||  || |  || |    |     |  |  |||||| ||| ||
Db   129 AAATTAGACATCTTCAT--TTAGAGAGAAGATGGAAACAACATTGCTTTTCTTTTTCT 186

Qy   161 CAATTAAATATGTGAATCAAAAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGG 220
          ||| |||||||||| |||||||| |||| |||||||||| ||| ||| |||||||||
Db   187 CAAATAAATATGTGTGAATCAAAAGAAAAGAAAACTTTTTTCAAGTAATACATGGGTCAGGA 246

Qy   221 AAAGGAGAAACAAGCATCGAGGCCAAAATCAGAGAGCGAAAGAAAAAGGATAGACTAAGT 280
          |||| ||||||||| | |   |||||||||||||| ||||  |||||| || |||||||
Db   247 AAAGAAGAAACAACAAGCAAAGAAGCCAAAATCAGAGCTAGGAGAAAAGAAATAGAACTAAGT 306

Qy   281 CTTCTCCTACGAGGAGCCTGACTTCCATGGAGAGACTCAAGCCAGTAGATCTCGCCCCTCTTG 340
          |||||  |  |  ||| ||||| || ||| | | || || || ||||||||| |  ||| ||
Db   307 CTTCTGTGCAGAAACCTGAGTTTCATGAAGACACACCCGCTCCAGTAGATCTTGGGCACTTG 366

Qy   341 GCCAAAGAAACAAGAGTCTCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAA 400
          |||||||||||||||| |||| ||||||| ||||||||||||| |||| ||||||||||
Db   367 GCCAAAGAAACAAGAGAGTCTCCCCTGAAGAGCAGTGAAATGGGGTGAATCATTTGACAAA 426

Qy   401 TTGCTCTCCTCATAGAGATGGAGTGGGATGCTTTTACCAGATTTCTAAAACTGAATTCAGT 460
          |||| | ||| ||||||||| |   | |  ||||||||||||||||||||||||| ||||
Db   427 CTGCTTTCCCATAGAGATGGGACTAGAGAGCGTTTTACCAGATTTCTAAAACTGAATTCAGT 486

Qy   461 GAGGAGACACATTGAATTTGGGTCGCCTGTGAAGACTTCAAGGAATGCAAGGAACCTCAA 520
          || || | | ||||||||||| | ||||||||| ||||||| || |||||| |||||||
Db   487 GAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGACCTCAA 546

Qy   521 CAAATCATCCTAAAAGCCAAAGGCAATCTGAGAAATTCATTCAGAATGATGCCCCCAAA 580
          |||| || ||||||| ||||||||| |  |||| ||| ||| ||| ||| || ||||
Db   547 CAAATTCACCCTAAAGCAAAAGCAATATATGAGAGAAATTTATACAGACTGAGTCCCCAAAA 606

Qy   581 GAGGTAACATTGATTTTCATACTAAGAGAGCATGCTAGAGCATCGCCCAGCCCACT 640
          |||||||| ||||||||||  | ||| ||  || |  |||||| | | || || |
Db   607 GAGGTAACCTTGATTTTCCACAAAGAAGTCATTACAACAGCATCACTCACCTACC 666

Qy   641 CTCCACAGTTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTAT 700
          |||||||||||||| | | ||||||||||||||| | ||||||||||||||||||||||
Db   667 CTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACATGACAGTTAT 726

Qy   701 AAACGCTTTTTGAAATCTGACACCTACTTACATTTGATAGGAAGAACCTCAGAGACCA 760
          | ||| |||| |||||| |||| |||| || |||||| ||  ||||||||||||||||
Db   727 ACACGTTTTCTGAAATCGACATCTATTTAGACTTGATGGGAGGACCTCAGAGACCA 786

Qy   761 ACAAACCTTAGGAGACGACGATCACGACGCATCACGATCATTCATTTTTACTTACTTTCCAAGATGTAAAGTCA 820
          ||||| |||||||||||||||||||| | |||| ||||||||||  ||| | || |||||||||||| || ||
Db   787 ACAAATCTTAGGAGACGACGATCACGCTCATTCATTTACCTGCAATGAATTCCAAGATGTACAATCA 846

Qy   821 GATGTTGCCATTTGGTTATGGAGTAAAAGTCATTTGTCTTCTTTGATAGTGTATGTGTAT 880
          ||||||||||||||||||||| ||||| | |||||| |||  ||| || |||| ||| ||
Db   847 GATGTTGCCATTTGGTTATGGTTATAAAGAAAATTGATTTTGCTCATTTTATGACAAACTTATAC 906

Qy   881 ATCTAAAAATATATACTAATAATGTACTTCTAAAATAATAGAGCTGTGTGTATAAGAGAGA 940
          |||
Db   907 ATC-------TGCTTCTCAACATATCCGATGTTTATGTTAAGA 941

Qy   941 GATGATTTCATTT-----TAAAATACACACCATGCAAATACATATTTAAAATGTAAGAACTT 994
          ||||| || ||||     |||||||||| | ||||||||||| |||||||||||||||| |
Db   942 TTTGGTCCCATCCTTTAAACTGAGAAATAGTCATGTGAAATTATTTTAAAAATGTAAAAAC 1001

Qy   995 TTTATATTATACTAAAATAATT-CATCATCTATCTTCCGAAATATTTTATGAAAATCTATC 1054
          | ||| | | || |||| ||||  || |||| ||| |||||| |||| |||||||||||||
Db  1002 AAAACTTTCTGCTAACAAAATCATACAGTATTCTGCCAGTATATCTGTAAAACCTTCTA 1061

Qy  1055 TGATAATTCTATTCTAATAAAATTC 1078
          ||  |  ||  ||||  |||||| ||
Db  1062 TTTGATGTGTCATTCCATTTATAATC 1085
```

RESULT 8
ABQ98617
ID   ABQ98617 standard; DNA; 848 BP.
XX
AC   ABQ98617;
XX
DT   04-NOV-2002 (first entry)
XX
DE   Human ORF424 coding sequence.
XX
KW   Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW   Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW   human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW   cancer; cardiovascular disease; allergy; autoimmune disease;
KW   wound healing; blood coagulation disorder; inflammatory disorder; ds.
XX
OS   Homo sapiens.
XX
PN   US2002082206-A1.
XX
PD   27-JUN-2002.
XX
PF   30-MAY-2001; 2001US-00867550.
XX
PR   30-MAY-2000; 2000US-0208427P.
XX
PA   (LEAC/) LEACH M D.
PA   (MEHR/) MEHRABAN F.
PA   (CONL/) CONLEY P B.
PA   (TOPP/) TOPPER J N.
PA   (LAWD/) LAW D.
XX
PI   Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
DR   WPI; 2002-626554/67.
DR   P-PSDB; ABP64054.
XX
PT   New polypeptide designated ORFX are present in human atherogenic cells
PT   and are useful to prevent and treat ORFX-associated disorders including
PT   cancer, allergy, wound healing or autoimmune, cardiovascular or
PT   inflammatory disease.
XX
PS   Claim 2; SEQ ID NO 847; 78pp; English.
XX
CC   The present invention relates to novel human ORFX polypeptides and their
CC   coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC   were discovered in human atherogenic cells, in particular in platelets
CC   and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC   many other tissues as well. Atherogenic cells are cells which have the
CC   potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC   nucleic acids are useful for treating or preventing a pathological
CC   condition associated with an ORFX-associated disorder, e.g. cancer,
CC   cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC   coagulation disorders or inflammatory disorders. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from the USPTO web site at
CC   seqdata.uspto.gov/sequence.html?DocID=2002082206
XX
SQ   Sequence 848 BP; 316 A; 149 C; 159 G; 223 T; 0 U; 1 Other;

```
Query Match    48.2%;    Score 560.6;  DB 6;  Length 848;
Best Local Similarity   82.0%;  Pred. No. 1.6e-113;
Matches   670;  Conservative   0;  Mismatches 144;  Indels   3;  Gaps   2;

Qy     1 TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGATAGCGCTTTATTC-AGGATGTTTTCC 59
          || ||||||| ||  |||| |  ||||||| |   || ||| |||| | || |||||
Db    34 TTCTTTTTGTAAACATTACTGTAAGAGTGTGATAACTTTTTTATTCTTACTACTATATG 93

Qy    60 TATGAAATAGCATTCATCGTGTGGGAGAGAGAAGGACTAAGGACTAAGGAATCTGACATCTGTTGGT 119
```

```
Db          |||  ||||  |||  ||||       ||  ||||           ||||  |||||||   |||||
         94 TATGGAATAGTATTAATAAATGAACTAGGGAAGGATGTAATAAATTAGACATCTTCAT 153

Qy      120 CACTGGGACAGAATATGGATATGTCACTGGTTTCTTCTCTCAATTAAATATGTGTGAAT 179
                                                                       ||||||
Db      154 T--TTAGAGGAGAGAAAACTTTTTCAAACTAATGCTTTCTTTTCTTTTCTCAAATAAATATGTGTGAAT 211

Qy      180 CAAAAGAGAGAAAACTTTTTCAAACTAATGCTCATGGGTCAGGGAAAGAAGAACAAGCATCG 239
            |||||||||||  ||||||||||||||||||||||||  ||||||||||||||||||
Db      212 CAAAAGAAAAACTTTTTTCAAGTTAATACATGGTTCAGGAAAGAAGAAACAAGCAAAG 271

Qy      240 AGGCCAAAATCAGAGCGCAAAGAACGAAAGAAAAAAGGAATAGACTAAGTCTTCCTACAGAGGCCTG 299
            ||||||||||||||||||||||||||||||||||||||    ||    || ||||||||||
Db      272 AAGCCAAAATCAGAGCTAAGGAAGAAATAAGAAATAGACTAAGTCTTCTGTGTCGAGAAACCTG 331

Qy      300 ACTTCCATGGAGAGAGACTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCT 359
            |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db      332 AGTTTCATGGAGAGACACACCCGCTCCAGTAGATCTCGGGCACTTGGCCAAAGAAACAAGAGTCT 391

Qy      360 CTCCTGAGAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGATG 419
            ||||||||||||||||  ||| ||||||||||||  || |||||||||||||||||
Db      392 CCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTTGACAAACTGCTTTCCCATAGAGATG 451

Qy      420 GAGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAAAATATTGAATTTT 479
            |||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db      452 GACTAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAAGAAAATATTGAATTTT 511

Qy      480 GGGTCCGCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCATCTAAAAGCAA 539
            ||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db      512 GGATAGCCGTGAAGATTTCAAGAAATGCAAGGGAACCTCAACAAATTCACCTTAAAGCAA 571

Qy      540 AGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGTTAACATTGATTTTC 599
            |||||||| || ||||||| ||||||||||||||  |||||||||||||||||||||||
Db      572 AAGCAATATATGAGAAATTCATTCAGAATTATACAGACTGATGCCCCAAAGAGTTAACTTGATTTTC 631

Qy      600 ATACTAAAGAGAGTAATTGCTAAGACGATCGCCCAGCCACTCTCCACAGTTTTGATACGG 659
            ||||||||||||||||||| || ||||||||||||||||||||||||||||||
Db      632 ACACAAAAGAAGTCATTACAAACAGCATCACTCAACCTCAACCTACCCCAGTTTTGATGCTG 691

Qy      660 CACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAAATCTG 719
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      692 CACAAAGCAGAGTGTATCAGTCATGCATGGAACAACAAGACAGTTATCATGGGTATTATACCGTTTTCTG 751

Qy      720 AGACCTACTTACATTTGATAGAAGAAGAGACCTCAGAGACCAACAACAAACCTTAGGACGACGAT 779
            ||  ||  |||  |  || |||||||||||||||||||||||||||||||||||||||
Db      752 ACATCATTTAGACTTGATGGAAGAAGAGACCTCAGAGACCAACAAATCTTAGGAGGACGAT 811

Qy      780 CACGATCATTTACTTACACAATGATTTCCAAGATGTAAA 816
            |||||||||||  ||| |||||||||||||||||||||
Db      812 CACGCTCATTTACCTGCAATGAATTCCAGATGTACA 848
```

RESULT 9
AAS18331
```
ID   AAS18331 standard; cDNA; 1486 BP.
XX
AC   AAS18331;
XX
DT   12-MAR-2002 (first entry)
XX
DE   cDNA from a human thyroid library encoding C-terminal region of RGS18.
XX
KW   Human; regulator of G protein signalling; RGS18; arterial thrombosis;
KW   platelet activation dysfunction; myocardial infarction; stroke;
KW   coronary artery disease; cerebrovascular disease; unstable angina;
KW   deep vein thrombosis; systemic thromboembolism; anti-coagulant;
KW   invasive cardiac procedure; thyroid; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             3..212
FT                   /tag= a
```

```
PT   Isolated regulator of G protein signaling polypeptide, useful for
PT   prevention/treatment of platelet activation dysfunction, such as arterial
PT   thrombosis, myocardial infarction, coronary artery disease and stroke.
XX
PS   Example 2; Page 113-114; 127pp; English.
XX
CC   The present invention relates to the isolation of polynucleotide
CC   sequences, that encode a novel regulator of G protein signalling (RGS)
CC   polypeptide, RGS18, from human platelets. The invention also provides
CC   nucleotide primers and probes specific for an RGS18 nucleic acid. The
CC   sequences of the invention are useful for the manufacture of a medicament
CC   for the prevention or treatment of a platelet activation dysfunction,
CC   such as arterial thrombosis, myocardial infarction, coronary artery
CC   disease, stroke, cerebrovascular disease, unstable angina, deep vein
CC   thrombosis, systemic thromboembolism, and in invasive cardiac procedures
CC   for anti-coagulant purposes. The sequences of the invention can also be
CC   used for the detection of RGS18 nucleic acids and RGS18 polypeptides. The
CC   present sequence represents a cDNA sequence from an Incyte clone derived
CC   from a human thyroid library. The sequence is used to isolate the full
CC   length cDNA encoding RGS18 from human platelets
XX
SQ   Sequence 1486 BP; 529 A; 235 C; 240 G; 482 T; 0 U; 0 Other;
```

```
Query Match             16.5%;  Score 192.2;  DB 6;  Length 1486;
Best Local Similarity   68.4%;  Pred. No. 1.6e-32;
Matches   311;  Conservative   0;  Mismatches 113;  Indels  31;  Gaps  2;
```

```
Qy      630 CCCAGCCACTCTCCACAGTTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGAAC 689
            |||||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Db        1 CTCAACCTACCCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGTCATGGAAC 60

Qy      690 ATGACAGTTATAAACGCTTTTTGAAATCTGAATCTGACATCTTACATTTGATAGAAGAAGAGAC 749
            |||||||||  ||||| ||||||||||||||||  ||||||||||||||||||||||||||
Db       61 AAGACAGTTATCATGGGTATTATACCGTTTTCTGAATCTGACATCTTAGAGTTGATGGAAGAAGAGAC 120

Qy      750 CTCAGAGACCAACAAACCTTAGGAGGACGACGATCACGATCATTTACTTACTTCATTTTGCCAAG 809
            ||||||||||||||||  |||||||||||||||||||||||||||| ||||||  ||||||
Db      121 CTCAGAGACCAACAAATCTTAGGAGGACGACGATCACGCTCATTTACCTGCAATGAATTCCAAG 180

Qy      810 ATGTAAAGTCAGATGTGCCATTGGTATGAGTAAAGTCATTTGTTGATATG 869
            ||||||||||||||| |||||||||||||  ||||| ||||||| |||||
Db      181 ATGTACAATCAGATGTGCCATTGGTTATAAAGAAAATTGATTTTGCTCATTTTTATGA 240

Qy      870 TGTATGTGTATATCTAAAATATATCTAAAATACTAATACTAATGTGTACTTGTACTTGT 929
            |||||||||||||||     ||||||||| ||||||||||||| --------------
Db      241 CAAACTTATACATC----------------------------TGCTTCTAACATCGCATGT 275

Qy      930 GTATAAGAAGAGATGATTTCATTTT----------TAAAATACACCATGCAAATACATATTAAA 983
            ||||  ||| ||||||| |||||||         ||||||||  ||||||||||||||||||
Db      276 TTATGTTAAGATTTGGTCCCATCCTTAAACTGAATAATTCATCATCTATCTTCCGAAATATTTTAT 335

Qy      984 TGTAAGAACTTTTTATTATTATTATTACTAAAATAATTCATCATCTATCTTCCGAAATATTTTAT 1043
            |||   ||||||||||||||  |||| | |||||||||||||||||||||||||||||||||
```

```
Db   336 AATGTAAAAACAAAACTTTCTGCTAACAAAATACATACAGTATCTGCCAGTATATTCTGT 395

Qy  1044 GAAAATCATCTGATATCTCATTCTAATAAAATTC 1078
         ||||| ||||||||||||||||||||||| ||||
Db   396 AAAACCTTCTATTTGATGTCATTCCATTTATAATC 430
```

RESULT 10
AAS18334
ID   AAS18334 standard; cDNA; 241 BP.
XX
AC   AAS18334;
XX
DT   12-MAR-2002 (first entry)
XX
DE   cDNA encoding partial human platelet RGS domain.
XX
KW   Human; regulator of G protein signalling; RGS18; arterial thrombosis;
KW   platelet activation dysfunction; myocardial infarction; stroke;
KW   coronary artery disease; cerebrovascular disease; unstable angina;
KW   deep vein thrombosis; systemic thromboembolism; anti-coagulant;
KW   invasive cardiac procedure; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..240
FT   /tag= a
FT   /partial
FT   /product="Partial RGS domain"
FT   /note="This sequence encodes for residues 2-81 of
FT   AAU10747 and lacks both start and stop codons"
XX
PN   WO200183514-A2.
XX
PD   08-NOV-2001.
XX
PF   26-APR-2001; 2001WO-US013540.
XX
PR   28-APR-2000; 2000US-0200786P.
PR   02-AUG-2000; 2000GB-00018833.
XX
PA   (AVET) AVENTIS PHARM PROD INC.
XX
PI   Murray DL, Gagnon AW;
XX
DR   WPI; 2002-055453/07.
DR   P-PSDB; AAU10747.
XX
PT   Isolated regulator of G protein signalling polypeptide, useful for
PT   prevention/treatment of platelet activation dysfunction, such as arterial
PT   thrombosis, myocardial infarction, coronary artery disease and stroke.
XX
PS   Claim 1; Page 115; 127pp; English.
XX
CC   The present invention relates to the isolation of polynucleotide
CC   sequences, that encode a novel regulator of G protein signalling (RGS)
CC   polypeptide, RGS18, from human platelets. The invention also provides
CC   nucleotide primers and probes specific for an RGS18 nucleic acid. The
CC   sequences of the invention are useful for the manufacture of a medicament
CC   for the prevention or treatment of a platelet activation dysfunction,
CC   such as arterial thrombosis, myocardial infarction, coronary artery
CC   disease, stroke, cerebrovascular disease, unstable angina, deep vein
CC   thrombosis, systemic thromboembolism, and in invasive cardiac procedures
CC   for anti-coagulant purposes. The sequences of the invention can also be
CC   used for the detection of RGS18 nucleic acids and RGS18 polypeptides. The
CC   present cDNA sequence encodes for a partial human platelet RGS domain
XX
SQ   Sequence 241 BP; 92 A; 52 C; 44 G; 53 T; 0 U; 0 Other;

Query Match   15.3%;   Score 177.6; DB 6; Length 241;
Best Local Similarity 83.8%; Pred. No. 1.7e-29;
Matches 201; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```
Qy  461 GAGGAGAGAACATTGATTTTGGGTCGCTCGTGAGACTTCAAGAGAATGCAAGGAGAACCTCAA 520
        |||||| ||||||||||||||||| |||||||||||||||||||||||||||| |||||||||
Db    1 GAGGAAAATCGGAGTTCTGGATAGCCTGTGAAGATTCAAGAGAAAGCAAGGGACCTCAA 60

Qy  521 CAAATCATCCTAAAGCAAAGGCAATCATCATGAGAAATTCATTCAGAATGAATGCCCCCAAA 580
        |||||| |||||||||||| ||| ||||||| |||||||||| |||||| ||| |||||||
Db   61 CAAATCACCTTAAAGCAAAGCAAAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAA 120

Qy  581 GAGGTTAACATTGATTTTCATACTAAGAAGTAATTGCTAAGAGCATCCCCAGCCCACT 640
        |||||||||| ||||||||| ||| |||||||| |||  ||||| ||| |||| ||||
Db  121 GAGGTTAACCTTGATTTTCACACAAAGAGTCATTACAAACAGCATCACCTCAACCTACC 180

Qy  641 CTCCACAGTTTTGATACGGACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTAT 700
        ||||||||||||||| | | ||||||||| ||||| |||||||||||||  ||||| ||||
Db  181 CTCCACAGTTTTGATGCTGCACAAAGCAGCAGAGTGTATCAGCTCATGGAAAACGACAGCTAT 240
```

RESULT 11
ADD34052
ID   ADD34052 standard; DNA; 681 BP.
XX
AC   ADD34052;
XX
DT   15-JAN-2004 (first entry)
XX
DE   Mouse mitochondrial DNA sequence SEQ ID NO:1828.
XX
KW   ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW   mitochondrial disease; oxidative phosphorylation dysfunction;
KW   oxidative stress; apoptosis; aging.
XX
OS   Mus musculus.
XX
PN   WO200302020-A2.
XX
PD   13-MAR-2003.
XX
PF   30-AUG-2002; 2002WO-US027886.
XX
PR   30-AUG-2001; 2001US-0316323P.
PR   31-AUG-2001; 2001CA-02356540.
XX
PA   (UYEM-) UNIV EMORY.
XX
PI   Wallace DC, Levy S, Kerstann K, Procaccio V;
XX
DR   WPI; 2003-300821/29.
XX
PT   Array containing probes for genes involved in mitochondrial biology,
PT   useful for determining mitochondrial biology gene expression profiles for
PT   use in diagnosing pathologies and identifying biochemical pathways.
XX
PS   Claim 2; SEQ ID NO 1828; 201pp; English.
XX
CC   The invention relates to a novel array comprising at least two isolated
CC   nucleotide molecules, each molecule having a sequence capable of uniquely
CC   hybridising to a nucleic acid molecule which is an expression product of
CC   a gene involved in mitochondrial biology. The array comprises two or more
CC   isolated nucleic acid molecules or spots, each molecule having a sequence
CC   chosen from sequence of 994 human probes and 2046 mouse probes. An array
CC   of the invention is useful for determining an expression profile of a
CC   mouse or human sample containing nucleic acid, by contacting the array
CC   with the sample under conditions allowing selective hybridisation, and
CC   measuring hybridisation of nucleic acid in the sample to the array to
CC   produce an expression profile. The array is also useful for determining
CC   an expression profile of a first labelled sample containing nucleic acid
CC   relative to a second, differently labelled sample containing nucleic
CC   acid. The second sample is a reference or a standard. An array is useful
CC   for determining an expression profile diagnostic of an energy-metabolism-
CC   related physiological condition. An array of the invention is useful for
CC   determining mitochondrial biology gene expression profiles of organisms,
CC   such as human, mice and closely related species, tissue and organs of

CC   such organisms, which are useful for determining expression profiles
CC   diagnostic of energy metabolism-related physiological conditions,
CC   diagnosing such physiological conditions, identifying biochemical
CC   pathways, genes, and mutations involved in such physiological conditions,
CC   identifying therapeutic agents useful for preventing and/or treating such
CC   physiological conditions, evaluating and/or monitoring the efficacy of human
CC   such therapies, and creating and identifying animal models of human
CC   energy metabolism-related physiological conditions. An array is also
CC   useful for defining expression signatures or profiles for mitochondrial
CC   diseases, as well as distinguishing clinical disorders that result from
CC   oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
CC   apoptosis and aging. An array of the invention contains probes of genes
CC   not previously recognised to participate in mitochondrial biology. The
CC   sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA
CC   clones used to make the probes of the invention. Some sequences are not
CC   present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
CC   1906, 2408 and 2643.
XX
SQ   Sequence 681 BP; 217 A; 164 C; 157 G; 143 T; 0 U; 0 Other;

Query Match   13.4%; Score 156.2; DB 9; Length 681;
Best Local Similarity   64.0%; Pred. No. 1.1e-24;
Matches   236;   Conservative   0; Mismatches 133; Indels   0; Gaps   0;

Qy   359 TCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTGCAAATTGCTCTCTCATAGAGAT 418
Db   220 TCTCCTGAGGAAGCCGAGCTCTCGGGCCAGAAGCATTTGATGAACTGCTGGCCAGTAAATAT 279

Qy   419 GGAGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 478
Db   280 GGGCTGCTTGCTGCATTCAGGGCGTTTTTAAAGTCCGAGTTCTGTGAAGAAAACATTGAATTC 339

Qy   479 TGGGTCGCCTGTGAAGAACTTCAAGAAGACTTCAAGGACCTCAACAAATCATCCTAAAAGCA 538
Db   340 TGGTTGGCTTGTGAAGACTTCAAGAAGACTTCAAAAACCAAATCACCCCAAAAACTGTCCTCAAAAGCA 399

Qy   539 AAGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTT 598
Db   400 AGGAAAATTCTATACCGACTTCATAGAGGAAGGAAGCTCCAAAGAGATAAACATAGACTTC 459

Qy   599 CATACTAAAGAAGTAATTGCTAAGAGACATCGCGCCCAGCCCACTTCTCCACAGTTTGATACG 658
Db   460 CAAACGAAAATCTCTGATTGCCCAAAATATCCAAGAGGCTACAAGTGGCTGCTTCACCACA 519

Qy   659 GCACAAAGCAGAGTGTACCAGCTCATGAACATGACAGTTATAAACGCTTTTGAAATCT 718
Db   520 GTTCAGAGAGAGGGTGTACAGTTTGATGAGAACAATTCTTATCCTCGGTTCTTGGAGTCC 579

Qy   719 GAGACCTAC 727
Db   580 GAATTCTAC 588

RESULT 12
ABT10881
ID   ABT10881 standard; cDNA; 1345 BP.
XX
AC   ABT10881;
XX
DT   04-DEC-2002 (first entry)
XX
DE   Human breast cancer associated coding sequence SEQ ID NO: 1015.
XX
KW   Human; breast specific gene; breast cancer; differential expression;
KW   cytostatic; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200259271-A2.
XX
PD   01-AUG-2002.
XX
PF   25-JAN-2002; 2002WO-US002176.

XX
PR   25-JAN-2001; 2001US-0263757P.
PR   25-APR-2001; 2001US-0286090P.
PR   23-MAY-2001; 2001US-0292517P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Orr MS, Nation M, Diggans JC, Zeng W;
XX
DR   WPI; 2002-674803/72.
XX
PT   Diagnosing breast cancer in a patient comprises detecting the level of
PT   gene expression in cell or tissue samples, where a differential gene
PT   expression is indicative of breast cancer.
XX
PS   Claim 1; SEQ ID NO 1015; 260pp + Sequence Listing; English.
XX
CC   The present invention relates to methods of diagnosing breast cancer in a
CC   patient, which comprise detecting the level of expression in a tissue
CC   sample of two or more genes selected from those shown in ABT09867-
CC   ABT11112, where a differential expression of the genes indicates breast
CC   cancer. The methods are useful in diagnosing, treating, detecting the
CC   progression, and in monitoring treatment of breast cancer in patients.
CC   The methods are also useful as a screening tool for agents that modulate
CC   the onset or progression of breast cancer. The breast cancer genes may be
CC   used as diagnostic markers for the prediction or identification of the
CC   malignant state of breast tissue, for confirming the type and progression
CC   of cancer, and for drug screening and assays. The present sequence is a
CC   coding sequence of the invention. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub.published_pct_sequences
XX
SQ   Sequence 1345 BP; 443 A; 254 C; 304 G; 344 T; 0 U; 0 Other;

Query Match   13.1%; Score 152.4; DB 6; Length 1345;
Best Local Similarity   62.7%; Pred. No. 8.7e-24;
Matches   237;   Conservative   0; Mismatches 141; Indels   0; Gaps   0;

Qy   359 TCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTGCAAATTGCTCTCTCATAGAGAT 418
Db   249 TCTCCTGAGGAAGCCACAGCTGTGGTCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy   419 GGAGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 478
Db   309 GGTCTTGCTGCATTCAGGGCGTTTTTAAAGTCGGAATTCTGTGAAGAAAAATATTGAATTC 368

Qy   479 TGGGTCGCCTGTGAAGAACTTCAAGAAATGCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCA 538
Db   369 TGGCTGCCTGCTGGTGAAGACTTCAAAAAACCAAATCACCCCAAAAAGCGTGTCCTCAAAAGCA 428

Qy   539 AAGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTT 598
Db   429 AGGAAAATATATACTGACTTCATAGAGAAAGAGAAGCTCCAAAAGAGATAAACATAGATTTT 488

Qy   599 CATACTAAAGAAGTAATTGCTAAGAGACATCGCGCCCAGCCCACTTCTCCACAGTTTGATACG 658
Db   489 CAAACCAAAACTCTGATTGCCCAGAATATACAAGAGAGTACAAGTGGCTGCTTACACT 548

Qy   659 GCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAAATCT 718
Db   549 GCCCAGAAAAGGGTATACAGCTTGATGGAGAACAACTCTTATCCTCGGTTCTTGGAGTCA 608

Qy   719 GAGACCTACTTACCATTTG 736
Db   609 GAATTCTACCAGGACTTG 626

RESULT 13
ABK83834
ID   ABK83834 standard; cDNA; 1345 BP.
XX
AC   ABK83834;

```
Query Match      13.1%;   Score 152.4;  DB 6;  Length 1345;
Best Local Similarity  62.7%;   Pred. No. 8.7e-24;
Matches  237;  Conservative   0;  Mismatches 141;  Indels   0;  Gaps   0;

Qy     359 TCTCCTGAAGAAGCAGTGAAATGGCTGAATCATTGACAAATTGCTCTCTCATAGAGAT 418
           ||||||||||||||||||||||||||||||||  |||||| |||||| |||||||||
Db     249 TCTCCTGAGGAAGCACAGCTGTGGTCAGAAGCATTGACGAGCTGCTAGCCAGCAAATAT 308

Qy     419 GGAGTGGATGCTTTTACCAGATTCTTAAAACTCAGTGAGGAGAACATTGAATT 478
           |||||  |||||  |||||||| |||||||||||  |||||||||| ||||||
Db     309 GGTCTTGCTGCATTCAGGGCCTTTTTTAAAGTCGGAATTCTGTGAAGAAAATATTGAATTC 368

Qy     479 TGGGTGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCA 538
           |||||||||||||||||||||||||||  ||| ||||||||||||||||||| |||||
Db     369 TGGCTGGCCTGTGAAGACTTCAAAAAAACCAAATCACCCCAAAAGCTGTCCTCAAAAGCA 428

Qy     539 AAGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTAACATTGATTTT 598
           |||| ||||||| ||| ||||| |||||||||  ||||| ||||||| |||||||||||
Db     429 AGGAAATATATACTGACTTCATAGAAAAGGAAGCTCAAAAAGAGATAAACATAGATTTT 488

Qy     599 CATACTAAAGAAGTAATTGCTAAGAGCATCATCGCCCAGCCCACTCTTCCCACAGTTTTGATACG 658
           ||||||| |||||||| |||||||||  || |||||||||| ||||||||||||||||||||
Db     489 CAACCAAAACTCTTGATTGCCCAGAATATACAAGAAGCTACACAGTGGTGCTTTTACAACT 548

Qy     659 GCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTGAAATCT 718
           |||||||||||||||| ||||||||| |||| |||||||||| ||||||||||||||||
Db     549 GCCCAGAAAGGGTATACAGCTTGATGGAGAACAACTCCTTATCCTCCGTTTCTTGGAGTCA 608

Qy     719 GAGACCTACTTACATTTG 736
           ||||| |||||||
Db     609 GAATTCTTACCAGGACTTG 626

RESULT 14
ACC46750
ID   ACC46750 standard; cDNA; 1345 BP.
XX
AC   ACC46750;
XX
DT   05-JUN-2003  (first entry)
XX
DE   Human COPD related protein encoding cDNA SEQ ID NO:1.
XX
KW   Human; chronic obstructive pulmonary disease; COPD; chronic lung disease;
KW   gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO20029127-A2.
XX
PD   05-DEC-2002.
XX
PF   28-MAY-2002; 2002WO-EP005835.
XX
PR   31-MAY-2001; 2001GB-00013266.
XX
PA   (FARB ) BAYER AG.
XX
PI   Oellers N, Gehrmann M, Kallabis H, Hall R, Schulze T, Kroegel C;
XX
DR   WPI; 2003-140492/13.
DR   P-PSDB; ABP96779.
XX
PT   Predicting, diagnosing or prognosing chronic lung disease, by detecting a
PT   chronic obstructive pulmonary disease (COPD) gene in a biological sample.
XX
PS   Claim 8; Page 73-74; 214pp; English.
XX
CC   The present invention describes a method for predicting, diagnosing or
CC   prognosing chronic lung disease by detecting a chronic obstructive
CC   pulmonary disease (COPD) gene related polynucleotide (see ACC46750 to
CC   ACC46777, which encode the COPD related proteins in ABP96779 to
```

```
XX
DT   14-AUG-2002  (first entry)
XX
DE   Human cDNA differentially expressed in granulocytic cells #405.
XX
KW   Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW   viral infection; parasitic infection; protozoal infection;
KW   fungal infection; sterile inflammatory disease; psoriasis;
KW   rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW   cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW   adult respiratory distress syndrome; inflammatory bowel disease;
KW   Crohn's disease; ulcerative colitis; periodontal disease;
KW   granulocyte activation; chronic inflammation; allergy.
XX
OS   Homo sapiens.
XX
PN   WO20228999-A2.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US030821.
PR   03-OCT-2000; 2000US-0237189P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR   WPI; 2002-435328/46.
XX
PT   Detecting granulocyte activation by detecting differential expression of
PT   genes associated with granulocyte activation, which serves as diagnostic
PT   markers that is useful for monitoring disease states and drug toxicity.
XX
PS   Claim 1; SEQ ID NO 405; 114pp; English.
XX
CC   The invention relates to detecting (M1) granulocyte (GC) activation
CC   (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC   DNA chip analysis as given in the specification, and comparing the
CC   expression level to an expression level in an unactivated GC, where
CC   differential expression of Gs is indicative of GCA. Also included are
CC   modulating (M2) GA by contacting GC with an agent that alters the
CC   expression of at least one gene in Gs; (2) screening (M3) for an agent
CC   capable of modulating GCA or an inflammation (especially chronic) in a
CC   tissue, an allergic response in a subject, exposure of a subject to a
CC   pathogen or sterile inflammatory disease using the gene expression
CC   profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC   tissue, an allergic response in a subject, exposure of a subject to a
CC   pathogen or sterile inflammatory disease, by detecting the level of
CC   expression in a sample of the tissue of gene(s) from Gs, where the level
CC   of expression of the gene is indicative of inflammation; (4) treating
CC   (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC   response in a subject, exposure of a subject to a pathogen or sterile
CC   inflammatory disease, by contacting a tissue having inflammation with an
CC   agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC   is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC   for screening an agent capable of modulating GCA preferably in an
CC   inflammation in a tissue; M4 is useful for detecting an inflammation
CC   (especially chronic) in a tissue, an allergic response in a subject,
CC   exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC   psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC   cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC   respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC   disease, ulcerative colitis, periodontal disease; also bacterial
CC   infection, viral infection, parasitic infection, protozoal infection,
CC   fungal infection, and M5 is useful for treating one of the above
CC   conditions. The present sequence represents a gene differentially
CC   expressed in granulocytes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1345 BP; 443 A; 254 C; 304 G; 344 T; 0 U; 0 Other;
```

XX
PT   Combination of polynucleotide probes, useful as array elements in a
PT   microarray for monitoring the expression of a number of target
PT   polynucleotides.
XX
PS   Claim 1; SEQ ID NO 1278; 65pp; English.
XX
CC   The invention relates to a combination which, comprises a number of
CC   polynucleotide probes comprising a sequence selected from one of the 1490
CC   sequences mentioned in the specification. The combination is useful as an
CC   array element in a microarray for monitoring the expression of a number
CC   of target polynucleotides. The microarray is particularly useful in the
CC   diagnosis and treatment of cancer and immunopathology and neuropathology.
CC   The microarray is useful in diagnostics and treatment regimens, drug
CC   discovery and development, toxicological and carcinogenicity studies,
CC   forensics and pharmacogenomics. The microarray is also useful for
CC   monitoring progression of diseases and for developing sophisticated
CC   profiles for the effects of currently available therapeutic drugs. The
CC   combination is also useful for purifying a subpopulation of mRNAs, cDNAs
CC   and genomic fragments and in research and diagnostic applications. The
CC   array can detect changes in expression in a large number of genes coding
CC   for different signaling pathway populations which can be used to diagnose
CC   various diseases including cancer e.g. adenocarcinoma and leukaemia,
CC   immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
CC   and Parkinson's disease. The present sequence represents a polynucleotide
CC   probe of the invention. Note: The sequence data for this patent did not
CC   form part of the printed specification but was obtained in electronic
CC   format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=06500938B1
XX
SQ   Sequence 1345 BP; 443 A; 254 C; 304 G; 344 T; 0 U; 0 Other;

Query Match   13.1%; Score 152.4; DB 7; Length 1345;
Best Local Similarity   62.7%; Pred. No. 8.7e-24;
Matches 237; Conservative   0; Mismatches 141; Indels   0; Gaps   0;

Qy   359 TCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTGACAAATTGCTCTCCTCATAGAGAT 418
           ||||||||||||||||||||||||||||| | || |||||||||| ||||| ||||||
Db   249 TCTCCTGAGGAGCACAGCTGTGGTCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy   419 GGAGTGGATGCGTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 478
           ||||| || || ||||| ||| ||||| | || |||| ||| ||| |||||||||| |
Db   309 GGTCTTGCTGCATTCAGGGCTTTTTTAAAGTCGGAATTCTGTGAAGAAAATATTGAATTC 368

Qy   479 TGGGTCGCCTGTGAAGAATGCAAGGAACCTCAACAAATCATCCTAAAAGCA 538
           ||||||||||||||||||||||| |||||||||| |||||||||||||| ||
Db   369 TGGCTGGCCTGTGAAGACTTCAAAAAAACCAAATCACCCCAAAAGCTGTCCTCAAAAGCA 428

Qy   539 AAGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAGCCACTCTCCACAGTTTTGATACG 598
           ||||||||| | |||| |||||||||||| || |||||||||||||||
Db   429 AGGAAAATATATACTGACTTCATAGAGAAAAGGAAGCTTCAAAAGAGATAAACATAGATTTT 488

Qy   599 CATACTAAAGAGAGTAATTGCTAAGAGCATCGCCCAGCCACTCTCCACAGTTTTGATACG 658
           ||||| ||| | || ||| ||||||
Db   489 CAAACCAAAACTCTGATTGCCCAGAATATACAAGAAGCTCCAAAAGAGAATAAACAAACT 548

Qy   659 GCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGGCTTTTTGAAATCT 718
           ||||||||||| ||| ||||| ||| ||| || ||||| |
Db   549 GCCCAGAAAAGGTATACAGCTTGATGGAGAACAACTCTTATCTCCTTCTCGGAGTCA 608

Qy   719 GAGACCTACTTACATTG 736
           ||||| || ||||||
Db   609 GAATTCTACCAGGACTTG 626

RESULT 15
ACA56680
ID   ACA56680 standard; cDNA; 1345 BP.
XX
AC   ACA56680;
XX
DT   06-JUN-2003 (first entry)
XX
DE   Signalling pathway polynucleotide probe SEQ ID NO 1278.
XX
KW   Probe; ss; array element; microarray; Parkinson's disease;
KW   signalling pathway population; cancer; adenocarcinoma; leukaemia;
KW   immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease.
XX
OS   Unidentified.
XX
PN   US6500938-B1.
XX
PD   31-DEC-2002.
XX
PF   30-JAN-1998;   98US-00016434.
PR   30-JAN-1998;   98US-00016434.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Au-Young J, Seilhamer JJ;
XX
DR   WPI; 2003-352189/33.

CC   ABP96806). The method is useful for predicting, diagnosing or prognosing
CC   chronic lung disease in a biological sample. The COPD genes and proteins
CC   encoded by them from the present invention (I) can be used for treating
CC   or preventing chronic lung disease in a mammal. (I) can be used in an
CC   animal model for determining the efficacy, toxicity, or side effects of
CC   treatment with (I), and determining the mechanism of action of (I).
CC   ACC46778 to ACC46903 represent COPD related PCR primers and probes used
CC   in an example from the present invention
XX
SQ   Sequence 1345 BP; 443 A; 254 C; 304 G; 344 T; 0 U; 0 Other;

Query Match   13.1%; Score 152.4; DB 7; Length 1345;
Best Local Similarity   62.7%; Pred. No. 8.7e-24;
Matches 237; Conservative   0; Mismatches 141; Indels   0; Gaps   0;

Qy   359 TCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTGACAAATTGCTCTCCTCATAGAGAT 418
           ||||||||||||||||||||||||||||| | || |||||||||| ||||| ||||||
Db   249 TCTCCTGAGGAGCACAGCTGTGGTCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy   419 GGAGTGGATGCGTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 478
           ||||| || || ||||| ||| ||||| | || |||| ||| ||| |||||||||| |
Db   309 GGTCTTGCTGCATTCAGGGCTTTTTTAAAGTCGGAATTCTGTGAAGAAAATATTGAATTC 368

Qy   479 TGGGTCGCCTGTGAAGAATGCAAGGAACCTCAACAAATCATCCTAAAAGCA 538
           ||||||||||||||||||||||| |||||||||| |||||||||||||| ||
Db   369 TGGCTGGCCTGTGAAGACTTCAAAAAAACCAAATCACCCCAAAAGCTGTCCTCAAAAGCA 428

Qy   539 AAGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAGCCACTCTCCACAGTTTTGATACG 598
           ||||||||| | |||| |||||||||||| || |||||||||||||||
Db   429 AGGAAAATATATACTGACTTCATAGAGAAAAGGAAGCTTCAAAAGAGATAAACATAGATTTT 488

Qy   599 CATACTAAAGAGAGTAATTGCTAAGAGCATCGCCCAGCCACTCTCCACAGTTTTGATACG 658
           ||||| ||| | || ||| ||||||
Db   489 CAAACCAAAACTCTGATTGCCCAGAATATACAAGAAGCTCCAAAAGAGAATAAACAAACT 548

Qy   659 GCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGGCTTTTTGAAATCT 718
           ||||||||||| ||| ||||| ||| ||| || ||||| |
Db   549 GCCCAGAAAAGGTATACAGCTTGATGGAGAACAACTCTTATCTCCTTCTCGGAGTCA 608

Qy   719 GAGACCTACTTACATTG 736
           ||||| || ||||||
Db   609 GAATTCTACCAGGACTTG 626

Search completed: August 20, 2004, 06:15:37
Job time : 487.176 secs

```
                  GenCore version 5.1.6

          Copyright  (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    August 20, 2004, 05:14:44 ; Search time 92.2662 Seconds
                                     (without alignments)
                                     7001.078 Million cell updates/sec

Title:           US-09-894-749-3
Perfect score:   1164
Sequence:        1 ttttgtaagaaaatctga........gtactagtggatccgagctc 1164

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:   682709 seqs, 277747446 residues

Total number of hits satisfying chosen parameters:     1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database  :  Issued_Patents_NA:*
             1: /cgn2_6/ptcdata/2/ina/5A_COMB.seq:*
             2: /cgn2_6/ptcdata/2/ina/5B_COMB.seq:*
             3: /cgn2_6/ptcdata/2/ina/6A_COMB.seq:*
             4: /cgn2_6/ptcdata/2/ina/6B_COMB.seq:*
             5: /cgn2_6/ptcdata/2/ina/PCTUS_COMB.seq:*
             6: /cgn2_6/ptcdata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the total score being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1164	100.0	1164	3	US-09-244-314-3	Sequence 3, Appli
2	1164	100.0	1164	4	US-09-498-959-3	Sequence 4, Appli
3	596	51.2	2217	4	US-09-244-314-1	Sequence 1, Appli
4	596	51.2	2217	4	US-09-498-959-1	Sequence 1, Appli
5	305.4	26.2	736	4	US-09-833-381-1010	Sequence 1010, Ap
6	152.4	13.1	1345	4	US-09-016-434-1278	Sequence 1278, Ap
7	152.4	13.1	1345	4	US-09-023-655-1243	Sequence 1243, Ap
8	142	12.2	744	2	US-08-748-483-2	Sequence 2, Appli
9	142	12.2	2190	4	US-09-023-655-795	Sequence 795, App
10	139	11.9	800	4	US-09-016-434-1091	Sequence 1091, Ap
11	136	11.7	3208	4	US-09-566-921-111	Sequence 111, App
12	133.6	11.4	2638	4	US-09-016-434-1090	Sequence 1090, Ap
13	129.8	11.2	1398	4	US-09-016-434-1290	Sequence 1290, Ap
14	129.8	11.2	1398	4	US-09-702-705-322	Sequence 322, App
15	129.8	11.2	1398	4	US-09-736-457-322	Sequence 322, App
16	129.8	11.2	1398	4	US-09-614-124B-322	Sequence 322, App
17	129.8	11.2	1398	4	US-09-671-325-322	Sequence 322, App
18	129.8	11.2	1398	4	US-09-589-184-322	Sequence 322, App
19	122.6	10.5	606	2	US-08-463-081B-27	Sequence 27, Appli
20	122.6	10.5	606	2	US-08-461-379A-27	Sequence 27, Appli
21	122.6	10.5	606	2	US-08-462-390B-27	Sequence 27, Appli
22	122.6	10.5	606	3	US-08-463-074B-27	Sequence 27, Appli
23	122.6	10.5	606	3	US-08-465-585C-27	Sequence 27, Appli
24	122.6	10.5	606	3	US-08-652-446-27	Sequence 27, Appli
25	122.6	10.5	746	3	US-08-870-815-3	Sequence 3, Appli
26	122.6	10.5	746	3	US-08-949-004-3	Sequence 3, Appli
27	122.6	10.5	2383	1	US-08-274-318-1	Sequence 1, Appli
28	122.6	10.5	2383	2	US-08-754-108-1	Sequence 1, Appli
29	122.6	10.5	2406	2	US-08-463-081B-1	Sequence 1, Appli
30	122.6	10.5	2406	2	US-08-461-379A-1	Sequence 1, Appli
31	122.6	10.5	2406	2	US-08-462-390B-1	Sequence 1, Appli
32	122.6	10.5	2406	2	US-08-463-074B-1	Sequence 1, Appli
33	122.6	10.5	2406	3	US-08-465-585C-1	Sequence 1, Appli
34	122.6	10.5	2406	3	US-08-652-446-1	Sequence 1, Appli
35	102.8	8.8	981	2	US-08-829-110-2	Sequence 2, Appli
36	102.8	8.8	981	4	US-09-016-434-240	Sequence 240, App
37	102.4	8.6	2272	3	US-08-726-228-1	Sequence 1, Appli
38	100.4	8.6	2272	3	US-08-870-815-1	Sequence 1, Appli
39	100.4	8.6	2272	3	US-08-949-004-1	Sequence 1, Appli
40	100.4	7.6	1225	3	US-09-702-705-42	Sequence 4, Appli
41	89	7.2	498	c	US-09-736-457-42	Sequence 42, Appl
42	84	7.2	498	c	US-09-614-124B-42	Sequence 42, Appl
43	84	7.2	498	c	US-09-671-325-42	Sequence 42, Appl
44	84	7.2	498	c	US-09-589-184-42	Sequence 42, Appl
45	84					Sequence 42, Appl

ALIGNMENTS

```
RESULT 1
US-09-244-314-3
; Sequence 3, Application US/09244314
; Patent No. 6274362
; GENERAL INFORMATION:
;  APPLICANT: Hodge, Martin R.
;  APPLICANT: Yowe, David
;  TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
;  FILE REFERENCE: 5800-19, 035800/17480
;  CURRENT APPLICATION NUMBER: US/09/244,314
;  CURRENT FILING DATE: 1999-02-04
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 3
;  LENGTH: 1164
;  TYPE: DNA
;  ORGANISM: Mus sp.
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (134)..(841)
US-09-244-314-3

Query Match               100.0%;  Score 1164;  DB 3;  Length 1164;
Best Local Similarity     100.0%;  Pred. No. 2.7e-285;
Matches 1164; Conservative     0;  Mismatches     0;  Indels     0;  Gaps     0;

Qy         1 TTTTTGTAAGAAAAATCTGAGAGAAGATTCGGGATAGCGCTTATTCAGGATGTTTTCCT 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 TTTTTGTAAGAAAAATCTGAGAGAAGATTCGGGATAGCGCTTATTCAGGATGTTTTCCT 60

Qy        61 ATGAAATAGCATTCATCTGTGGGAGAGAGGACTAAGGAAATCTGACATCTGTTGGTC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ATGAAATAGCATTCATCTGTGGGAGAGAGGACTAAGGAAATCTGACATCTGTTGGTC 120

Qy       121 ACTGGGACAGAATATGGATATGTCACTGGTTTTCTTCCTTCTTCCAATTAAATATGTGAATC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 ACTGGGACAGAATATGGATATGTCACTGGTTTTCTTCCTTCTTCCAATTAAATATGTGAATC 180

Qy       181 AAAAGAGAAACTTTTTCAAACTAATGCATGCATGCAAAGAGAAACAAGCATCGA 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 AAAAGAGAAACTTTTTCAAACTAATGCATGCATGCAAAGAGAAACAAGCATCGA 240

Qy       241 GGCCAAAATCAGAGCCAGTAGATCTGCCCCTCTTGGCCAAAGAAACAAGAGGCCTGA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GGCCAAAATCAGAGCCAGTAGATCTGCCCCTCTTGGCCAAAGAAACAAGAGGCCTGA 300

Qy       301 CTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCCTTGGCCAAAGAAACAAGAGTCTC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 CTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCCTTGGCCAAAGAAACAAGAGTCTC 360
```

```
;  SEQ ID NO 3
;  LENGTH: 1164
;  TYPE: DNA
;  ORGANISM: Mus sp.
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (134)...(841)
US-09-498-959-3

   Query Match        100.0%;   Score 1164;   DB 4;   Length 1164;
   Best Local Similarity  100.0%;   Pred. No. 2.7e-285;
   Matches 1164;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

Qy      1  TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGAATAGCGCTTATTCAGGATGTTTTCCT   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGAATAGCGCTTATTCAGGATGTTTTCCT   60

Qy     61  ATGAAATAGCATTCATCTGTGGGAGAGAGGACTAAGGAAATCTGACATCTGTTGGTC   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  ATGAAATAGCATTCATCTGTGGGAGAGAGGACTAAGGAAATCTGACATCTGTTGGTC   120

Qy    121  ACTGGGACAGAATATGGATATGTCACTGGTTTCTTCTCTCAATTAAATATGTGTGAATC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  ACTGGGACAGAATATGGATATGTCACTGGTTTCTTCTCTCAATTAAATATGTGTGAATC   180

Qy    181  AAAAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGGAAGAAGAAACAAGCATCGA   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  AAAAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGGAAGAAGAAACAAGCATCGA   240

Qy    241  GGCCAAAATCAGAGCGAAAGAAAAAAGGAATAGAATCATGGGTCTTCTTCCTACAGAGGCCTGA   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GGCCAAAATCAGAGCGAAAGAAAAAAGGAATAGAATCATGGGTCTTCTTCCTACAGAGGCCTGA   300

Qy    301  CTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCTC   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  CTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCTC   360

Qy    361  TCCTGAAGAAGCAGCAGTGAAATGGCTGAATCATTGACAAATTGCTCTCCATAGAGATGG   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TCCTGAAGAAGCAGCAGTGAAATGGCTGAATCATTGACAAATTGCTCTCCATAGAGATGG   420

Qy    421  AGTGGATGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  AGTGGATGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG   480

Qy    481  GGTCGCCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAA   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  GGTCGCCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAA   540

Qy    541  GGCAATCTATGAGAAATTCATTCAGAAGTTGAAGAGCATCGCCCCAGCCCACTCTCCACAGTTTTGATACGGC   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GGCAATCTATGAGAAATTCATTCAGAAGTTGAAGAGCATCGCCCCAGCCCACTCTCCACAGTTTTGATACGGC   600

Qy    601  ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAAATCTGA   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAAATCTGA   660

Qy    661  GACCTACTTACATTGATAGAAGGAAGACCTCAGAGACCAACAAACCTTAGGAGACGATC   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  GACCTACTTACATTGATAGAAGGAAGACCTCAGAGACCAACAAACCTTAGGAGACGATC   720

Qy    721  ACGATCATTTACTTACTACAATGATTTCCAAGATGTAAAGTCAGATGTTGCCATTTGGTTATG   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  ACGATCATTTACTTACTACAATGATTTCCAAGATGTAAAGTCAGATGTTGCCATTTGGTTATG   780

Qy    781  AGTAAAGTCATTGTCTTCTTTTGATAGTGTATGTGTATATGTATGTATTAATATACTAATA   840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  AGTAAAGTCATTGTCTTCTTTTGATAGTGTATGTGTATATGTATGTATTAATATACTAATA   840
```

```
Qy    361  TCCTGAAGAAGCAGCAGTGAAATGGGCTGAATCATTTGACAAATGCTCTCCATAGAGATGG   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TCCTGAAGAAGCAGCAGTGAAATGGGCTGAATCATTTGACAAATGCTCTCCATAGAGATGG   420

Qy    421  AGTGGATGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  AGTGGATGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG   480

Qy    481  GGTCGCCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAA   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  GGTCGCCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAA   540

Qy    541  GGCAATCTATGAGAAATTCATTCAGAAGTTGAAGAGCATCGCCCCAGCCCAATTCCAGCACACTGGCG   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GGCAATCTATGAGAAATTCATTCAGAAGTTGAAGAGCATCGCCCCAGCCCAATTCCAGCACACTGGCG   600

Qy    601  TACTAAAGAAGTAATTGCTAAGAGCATGCCCCAGCCCACTCTCCACAGTTTTGATACGGC   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  TACTAAAGAAGTAATTGCTAAGAGCATGCCCCAGCCCACTCTCCACAGTTTTGATACGGC   660

Qy    601  GACCTACTTACATTGATAGAAGGAAGACCTCAGAGACCAACAAACCTTAGGAGACGATC   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  GACCTACTTACATTGATAGAAGGAAGACCTCAGAGACCAACAAACCTTAGGAGACGATC   720

Qy    721  ACGATCATTTACTTACTACAATGATTTCCAAGATGTAAAGTCAGATGTTGCCATTTGGTTATG   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  ACGATCATTTACTTACTACAATGATTTCCAAGATGTAAAGTCAGATGTTGCCATTTGGTTATG   780

Qy    901  CTAAATGTGTACTTCTAAAATATAGCTTGTGTATAAGAAGAGATGATTTCATTTTTAAAAT   960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  CTAAATGTGTACTTCTAAAATATAGCTTGTGTATAAGAAGAGATGATTTCATTTTTAAAAT   960

Qy    961  ACACCATGCAAATACATATTAAATGTAAGAACTTTTTATATTATACTAAAATAATTCATC   1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  ACACCATGCAAATACATATTAAATGTAAGAACTTTTTATATTATACTAAAATAATTCATC   1020

Qy   1021  ATCTATCTTCCGAAATATTTTATGAAATAAACGCTTTTCTAATTCTAATAAAATTCTT   1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  ATCTATCTTCCGAAATATTTTATGAAATAAACGCTTTTCTAATTCTAATAAAATTCTT   1080

Qy   1081  TATTTCTACAATAACAGTCAGTAAGAAGAAGCTTTGAAGCCGAATTCCAGCACACTGGCG   1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  TATTTCTACAATAACAGTCAGTAAGAAGAAGCTTTGAAGCCGAATTCCAGCACACTGGCG   1140

Qy   1141  GCCGGTACTAGTGGATCCGAGCTC   1164
           ||||||||||||||||||||||||
Db   1141  GCCGGTACTAGTGGATCCGAGCTC   1164
```

RESULT 2
US-09-498-959-3
; Sequence 3, Application US/09498959
; Patent No. 6410240
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; APPLICANT: Yowe, David
; TITLE OF INVENTION: RGS-Containing Molecules and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 5800-19A
; CURRENT APPLICATION NUMBER: US/09/498,959
; CURRENT FILING DATE: 2000-02-04
; EARLIER APPLICATION NUMBER: 09/244,314
; EARLIER FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 3.0

```
Qy  427  CTGCTTTCCCATAGAGATGGACTACTAGAGGCTTTTACCAGATTTCTTAAAACTGATTCAGT  486
         ||||| |||||||||||||||   |  ||||| |||||||||||||||||||||||||||||||
Db  461  GAGGAGCAATTGAATTTGGGTGCCTGTGAAGACTTCAAGAAATGCAAGGAGGACCTCAA  520
Db  487  GAAGAAAATATTGAATTTGGATAGCCTGTGAAGAGTTCAAGAGAAAAGCAAGGGGACCTCAA  546

Qy  521  CAAATCATCCTAAAAGCAAGGCAATCTAATGAGAAATCATTCAGAAATGCATCAGAGAATGCCCCAAA  580
         ||||| |||| ||||||||  |||| ||||||||| | |||| || |    ||| |||| ||||||||
Db  547  CAAATTCACCTTAAAGCAAAGCAATATATGAGAAATTTATACAGACTAAGTGCCCCAAAA  606

Qy  581  GAGGTTAACATTGATTTTCATACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACT  640
         ||||||||| |||||||||| || |||||| | || | |  ||||||||| ||  |||||
Db  607  GAGGTTAACCTTGATTTTCACACAAAAGAGTCATTACAAACAGCATCCACTCAACCTACC  666

Qy  641  CTCCACAGTTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGAACTGACACAGTTAT  700
         |||| |||||||||||| | |||||||||||||||| |||||| ||||||| ||||||||||
Db  667  CTCCCACAGTTTTGATGCTGCGCACAAAGCAGAGTGTATCAGCTTGATGGAACAAGACAGTTAT  726

Qy  701  AAACGCTTTTTGAAATCTGAGACGTCACATTCACTTACACAGTGTAAAGTCA  760
         | ||| |||||||||||||| |||| ||||| ||| |||||| ||| |||||
Db  727  ACACGTTTTCTGAAATCTGAAATCTGATGCATCATCTGATTGAGTGAAGTGAAGGAAGAAGACCTCAGAGACCA  786

Qy  761  ACAAACCTTAGGAGACGATCACGATCCATCATTCATTCACTTACTACATGTAAAGTCA  820
         | |||| ||||||||| |||| |||| || |||||| | |  |    |  |||||||
Db  787  ACAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCAAGAGATGTACAATCA  846

Qy  821  GATGTGCCATTGGTTGATGAGTAAAAGTCATTGCTTCTCTTTGATAGTGTATGTGTAT  880
         ||||||||||||| || |||| |||||||| |||| || |||||||  |||  |  |||
Db  847  GATGTGCCATTTGGTTGTATAAAGAAAATTGATTTTTGCTCATTTTATGACAAACTTATAC  906

Qy  881  ATCTAAAATATATACTAATAACTATACTGTACTTCTAAAATATAGCTTGTGTATAAGAGA  940
                                   ||||||   |||| ||  |||||| | ||| ||| |
Db  907  ATC------------------------TGCTTCAACATATCGCATGTTTATGTTAAGA  941

Qy  941  GATGATTCATTT------TAAAATACACCATGCAAATACATATTAAATGTAAGAACTT  994
         |||| ||||||||      | |   |||| ||    ||||  || ||| | |||||| ||
Db  942  TTTGGGTCCATCCTTTAAACTGAAATATGTCATGTGAAATTATTTTTAAAAATGTAAAAAC  1001

Qy  995  TTTATATTACTAAAATAATTCATCATCATTCCGAAATATTTTATGAAAAATCTATC  1054
         |||||  |  |||||  |||| ||||| ||||| ||||  ||| |||||||  | ||
Db  1002 AAAACTTTCTGCTAACAAAATACATACAGTGATATCGCAGTATATTCGTGTAAAACCTTCTA  1061

Qy  1055 TGATATTCTATTCTAATAAAATTC  1078
         ||||||| || |||  | ||||||
Db  1062 TTTGATGTCATTCCATTTATATC  1085
```

```
Qy  901  CTAAATGTGTACTTCTAAAAATATAGCTTGTGTGTATAAGAGAGATGATTCATTTTAAAAT  960
         |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  901  CTAAATGTGTACTTCTAAAAATATAGCTTGTGTGTATAAGAGAGATGATTCATTTTAAAAT  960

Qy  961  ACACCATGCAAATACATATTAAATGTAAGAACTTTTTATATTATACTAAAATAATTCATC  1020
         ||||||||||||| |||||||||||||||||||||||||||| ||||||| |||||||||
Db  961  ACACCATGCAAATAACATATTAAATGTAAGAACTTTTTATATTATACTAAAATAATTCATC  1020

Qy  1021 ATCTATCTTCCGAAATATTTTATGAAAATGATATTCTTCTAATGAATTCTAATAAAATTCTT  1080
         |||||||||||||||||||||| ||||||| ||| | |||||||||||||||||||||||||
Db  1021 ATCTATCTTCCGAAATATTTTTATGAAAATGATATTCTTCTTATTCTAATAAAATTCTT  1080

Qy  1081 TATTTCTACAATAACAGTCAGTAAGGAAGAAGCTTTGAAGCCGAATTCCAGCACACTGGGCG  1140
         |||||||||||| |||||||||||||||||||||| ||||||||||||||||||||||||||
Db  1081 TATTTCTACAATAACAGTCAGTAAGAAGAAGAAGCTTTGAAGCCGAATTCCAGCACACTGGGCG  1140

Qy  1141 GCCGGTACTAGTGGATCCGAGCTC  1164
         ||||||||||||||||||||||||
Db  1141 GCCGGTACTAGTGGATCCGAGCTC  1164
```

```
Qy   41  TTTATTCAGGATGCTTTTCCTATGAAATAGCATTCATCTGTGGGAGAGGAGGACTAAGG  100
         ||| |||| | ||| ||  ||||| |||||| || ||||  ||  |  |||| |||||||
Db   69  TTATTCACTCATGTATATGTGGAATAGTATTAATAAATGAACTAGGGAGGATGTAAT  128

Qy  101  AAATCTGACATCGTTTGGTCACTGGGACAGAGAATATGGATATGTCACTGGTTTCTCTCT  160
         ||||  |||||| |||||||||| |||| | |||||||| ||||| ||||| ||||| ||
Db  129  AAATTAGACATCTCCTTCAT--TTTAGAGAGAGATGGAAGAACTTTTTCTTTTCTTTTCT  186

Qy  161  CAATTAAATATGTGTGAATCAAAAGAAGAAAAACTTTTTCAAACTAAACTAACATGGTCAGGG  220
         |||| ||||||||||||||||||||||||||||||||||||| | ||||    ||||||||||
Db  187  CAAATAAATATGTGTGAATCAAAAGAAGAAAAACTTTTTCAAGTTAATACATAGTTCAGGA  246

Qy  221  AAAGAAGAAACAAGCATCGAGCCAAATCAGAGGCGAAAGAAAAAAGAAATAGACTAAGT  280
         |||||||||||||||   |||||||||||||| |  ||||||||| |||||||||||||
Db  247  AAAGAAGAAACAAGCAAAGAAGAAGCCAAATCAGAGCTAAGGAGAAAAAAGAAATAGACTAAGT  306

Qy  281  CTTCTCCTACAGAGGCCTGACTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCCTTG  340
         ||||| | | ||||  |||| ||||||| ||||  |   |||||||||||||| ||||||
Db  307  CTTCTTGTGTGCAGAAACCTGAGTTTCATGAAGACCACCCGCCTCCAGTAGATCTCGGCACTTG  366

Qy  341  GCCAAAGAAACAAGAGTCTTCCTGAAGAAGAGCAGTGAAATGGGCTGAATCATTTGACAAA  400
         ||||||||||||||||||||||  |||||||||||||||||||| ||||||||||||||||
Db  367  GCCAAAGAAACAAGAGTCTTCCCCTGAAGAGGAGTGAAATGGGGTGAATCATTTGACAAA  426

Qy  401  TTGCTTCTCCATAGAGATGGGAGTGGATGCGTTTTTACCAGATTTCTTAAAACTGAATTCAGT  460
         |||||| ||||||||||||||  || ||||  |||||||||||||||||||||||| |||||||
```

Best Local Similarity   75.8%;   Pred. No. 1.9e-141;
Matches   791; Conservative   0; Mismatches 220;   Indels   33;   Gaps   3;

```
Qy     41 TTTATTCAGGATGTTTCCTATGAAATAGCATTCATCTGTGGGAGAGAGGACTAAGG 100
             ||||      ||||        ||||        ||    || ||||
Db     69 TTATTCTACTATGTATATGTATGTGGAATAGTATTAATAAATGAACTAGGGAAGGATGTTAAT 128

Qy    101 AAATCTGACATCTGTTGGTCACTGGACAGAATATGGATATGTCACTGGTTTCTTCTCT 160
             ||||                 ||||||              |||||  |||||| |
Db    129 AAATTAGACATCTCTTCAT--TTTAGAGAGAGAAGGATGGAAACAACATTGCTTTTTCTTTTCT 186

Qy    161 CAATTAAATATGTGAATCAAAAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGG 220
             |||||||||||||  ||||      |||||      || |||  |||||||
Db    187 CAATAAATATGTGAATCAAAAGAAAAAACTTTTTTCAGTTAATACATGGTTCAGGA 246

Qy    221 AAAGAAGAAACAAGCATCGAGGCAAATCAGAGCGAAAGAAAAAAGGAATAGACTAAGT 280
             |||||||||||||||       ||||||||  |||||||  |||||||||  ||||
Db    247 AAAGAAGAAACAAGCAAAATCAGAGCGAAAGAAAAAAGGAAAATAGACTAAGT 306

Qy    281 CTTCTCCTACAGAGGCCTGACTTCCATGGAGACTCAAGCCAGTAGATCTGCCCCTCTTG 340
             ||||         ||||||||| |||||||||          ||||| |||||
Db    307 CTTCTTGTGCAGAAACCTGAGTTTCATCGAAGACACCCGCTCCAGTAGATCTGGGCACTTG 366

Qy    341 GCCAAAGAAACAAGAGTCTCTCTGAGAAGCAGTGAATCATTGACAAA 400
             |||||||||||||||||||||||        |||||||||||||||||
Db    367 GCCAAAGAAACAAGAGTCTCCCCTGAGAGGCAGTGAAATGGGTGAATCATTTGACAAA 426

Qy    401 TTGCTCTCTCATAGAGATGGAGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGT 460
             ||||       ||||||||||||||       ||||||||||||||||||||
Db    427 CTGCTTTTTCCCATAGAGATGGACTAGAGCGTTTTGCACAAAGCTTTTTACCAGATTTCTTAAAACTGAATTCAGT 486

Qy    461 GAGGAGAACATTGAATTTGGGTCGCCTGTGAAGACTTCAAGGAACCTCAA 520
             ||||||   |||||||||||   ||||||||||||||||||||||
Db    487 GAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA 546

Qy    521 CAAATCATCCTAAAAGCAAATCATTCAGGATATGCCCCCAAA 580
             ||||||||||       |||||||||| |||||
Db    547 CAAATTCACCTTAAAGCAAATAATATACAGACTGATGCCCCCAAAA 606

Qy    581 GAGGTTAACATTGATTTTCATACTAAAGAAGTAATGCTAGGAGCATCGCCCACT 640
             |||||||||||||||| |||||||||||      ||  ||||||||
Db    607 GAGGTTAACCTGATTTTTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCAACCTACC 666

Qy    641 CTCCACAGTTTGATAGGGCACAAAGCAGAGTGTACCAGCCAGTCAGATCATGAC 700
             ||||| ||||||||||||   ||||||||||||||||||||||||||
Db    667 CTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC 726

Qy    701 AAACGCTTTTGAAATCGAGACCTACTACATTTGATAGAGGAAGACCTCAG 760
             |||||||  || ||||||||||||||||||||||||||||||
Db    727 ACACGTTTTCTGAAATCTGACATCTATTTAGACTTAGACTTGATGGAGGAAGACCTCAG 786

Qy    761 ACAAACCTTAGGAGGACGATCACGATCATTTACTTACTTACTAAGTGTAAAGTCA 820
             |||||||||||||||||||| ||||||     ||||||||||||||||
Db    787 ACAAATCTTAGGAGGACGATCACGCATCACTTCATTTACTTACTTAACTGCAATATCA 846

Qy    821 AAGTCAGATGTTGCCATTTGGTTATGA 880
             |||||||||||||||||||||||||
Db    847 GATGTTGCCATTTGGTTATAAAGAAAAATTGATTTTGCTCATTTTTATGACAAACTTATAC 906

Qy    881 ATCTAAAAATATATACTAATAACTAATGTGTACTTCTAAAAATAGCTTGTGTGTAAGAGAGA 940
             ||||||||||||||||||||||||        ||| ||||||
Db    907 ATC--------------------------------TGCTTCTAACATATCGCATGTTTATGTTTAAGA 941

Qy    941 GATGATTCATTTT----------TAAAATACACCATGCAAATACATATTAAATGTAAGAACTT 994
             ||  ||||||||          ||    ||      ||||      |||||||||||
Db    942 TTTGGTCCCATCCTTTAAACTGAAATATGTCATGTGAAATTATTTTAAAAATGTAAAAAC 1001

Qy    995 TTTATTATTACTAAAATATATAAATATTTATGAAAATCTATC 1054
             ||||||||||||     |||||||||||||||||||||||
Db   1002 AAACTTTCTGCTAACAAATACATACAGTATCTGCCAGTATATTCTGTAAAAACTTCTA 1061

Qy   1055 TGATAATTCTATTCTAATAAAATTC 1078
             || ||||    | | ||||||||
Db   1062 TTTGATGTGTCATTCCATTTATAATC 1085
```

RESULT 5
US-09-833-381-1010
; Sequence 1010, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1010
; LENGTH: 736
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(736)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1010

Query Match   26.2%; Score 305.4; DB 4; Length 736;
Best Local Similarity 86.8%; Pred. No. 4.5e-68;
Matches 336; Conservative   0; Mismatches 51; Indels   0;   Gaps   0;

```
Qy    455 TTCAGTGAGGAGAACATTGAATTTGGGTCGCCTGTGAAGACTTCAAGGAAATGCAAGGAA 514
             ||||||||||||||||||||||||||    ||||||||||||  ||||||||||
Db      4 TTCAGTGAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGGA 63

Qy    515 CCTCAACAAATCATCCTAAAAGCAAATCATTCAGAATGATGCC 574
             ||||||||||||||        ||||||||||||  ||||||||
Db     64 CCTCAACAAATTCACCTTAAAGCAAATAATATACAGACTGATGCC 123

Qy    575 CCCAAAGAGGTTAACATTGATTTTCATACTAAAGAAGTAAGAAGCATCGCCCAG 634
             |||||||||||||||||||||| |||||||||||      ||  |||||||
Db    124 CCAAAGAGGTTAACCTGATTTTTGATTTTCACACAAAAGAAGTCATTACAAACAGCATCACTCAA 183

Qy    635 CCCACTCTCCACAGTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGAACATGAC 694
             |||||||||||| ||||||||   ||||||||||||||||||||||||||||||
Db    184 CCTACCCTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGAC 243

Qy    695 AGTTATAAACGCTTTTGAAATCGAGACCTACTACATTTGATAGAGGAAGAGACCTCAG 754
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    244 AGTTATACACGTTTTCTGAAATCTGACATCTATTTAGACTTGATGGAGGAAGAGACCTCAG 303

Qy    755 AGACCAACAAACCTTAGGAGGACGATCACGATCATTTACTTACTTACTAAGATGTA 814
             |||| |||||||||||||||||||| ||||||     ||||||||||||||||
Db    304 AGACCAACAAATCTTAGGAGGACGATCACGCATCACTTCATTTACTTACTTAACCTGCAATGATGTA 363

Qy    815 AAGTCAGATGTTGCCATTTGGTTATGA 841
             |||||||||||||||||||||||||
Db    364 CAATCAGATGTTGCCATTTGGTTATAA 390
```

RESULT 6
US-09-016-434-1278
; Sequence 1278, Application US/09016434
; Patent No.6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE

APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1243:
SEQUENCE CHARACTERISTICS:
LENGTH: 1345 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g292054
US-09-023-655-1243

Query Match   13.1%;   Score 152.4; DB 4;   Length 1345;
Best Local Similarity 62.7%;   Pred. No. 3e-29;
Matches 237;   Conservative   0; Mismatches 141;   Indels   0;   Gaps   0;

Qy 359 TCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGAT 418
Db 249 TCTCCTGAGGAAGCACAGCGTGTGTCAGAAGCATTTGACGAGCGTGCTGAGCACAATAT 308
Qy 419 GGAGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 478
Db 309 GGTCTTGCTGCATTCAGGGCTTTTTAAAGTCGGAATTCTGTGAAGAAAATATTGAATTC 368
Qy 479 TGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCA 538
Db 369 TGGCTGGCCTGTGAAGACTTCAAAAAAACCAAATCACCCCAAAAGTGTCCTCAAAAGCA 428
Qy 539 AAGGCAATCTATGAGAAATTCATTCAGAATGGGTTAACATTGATTTT 598
Db 429 AGGAAATATATACTGACTTCATAGAAAAGGAAGCTCCAAAAGAGATAAACATAGATTTT 488
Qy 599 CATACTAAAGAAGTAAGTAATTGCTAAGAGCATCGGCCCCAGCCCACTTCCACAGTTTTGATACG 658
Db 489 CAAACCAAAACCTGATTGCCCAGAATATACAAGAAGCTACAAGTGGCTGCTTTACAACT 548
Qy 659 GCACAAAGCAGAGTGTACCAGCTCATGGAACATGACGTTATAAACCGCTTTTGAAATCT 718
Db 549 GCCCAGAAAAGGGTATACAGCTTGATGGAGAACAACTTCTTATCCTCGTTCGTTCTTGGAGTCA 608
Qy 719 GAGACCTACTTACATTTG 736

RESULT 7
US-09-023-655-1243
; Patent No. 6607879
; Sequence 1243, Application US/09023655
; GENERAL INFORMATION:

```
Db     450  GAGGTGAATATTGACCACTTCACTAAGGACATCACAATGAAGAACCTGGTGGAACCTTCC  509
                                                                 
Qy     641  CTCCACAGTTTGATCGGCACAAAGCAGAGTGTACCAGCTCATCGAACATGACAGTTAT  700
                  ||||    ||   |||||   |||||       |||||  ||||||||||||
Db     510  CTGAGCAGCTTGCACATGGCCCAGAAAAGAATCCATGCCCTGATGGAAAAGGATTCTCTG  569

Qy     701  AAACGCTTTTGAAATCTGAGACCTACTTACATTTGAT  738
                  ||||||  ||||| || |||||||| ||||
Db     570  CCTCGCTTTGTGCGCTCTGAGTTTTATCAGGAGTAAT  607
```

RESULT 9
US-09-023-655-795
; Sequence 795, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/023,655
;   FILING DATE: HEREWITH
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 795:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 2190 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   LIBRARY: SYNORAT03
;   CLONE: 696878
US-09-023-655-795

```
Query Match        12.2%;    Score 142;  DB 4;  Length 2190;
Best Local Similarity  59.8%;  Pred. No. 1.5e-26;
Matches  238;  Conservative  0;  Mismatches 160;  Indels  0;  Gaps  0;

Qy     341  GCCAAAGAAACAAGAGTCTCTCCTGAAGAAGCAGTGAATCATTGACAAA  400
                 |||||||||||     ||  ||||||||||||||||||||||||||||||
Db     210  GCCAAGACCCAGAAAACCTCGCTGGACGAGGCCCTGCAGTGGCGGTGATTCCCTGGACAAA  269

Qy     401  TTGCTCTCTCATAGAGATGGAGATGGATGCTTTTAAAACTGAATTCAGT  460
                 |||||     |||||||  ||||| || |||||  ||||| |||||
Db     401  TTGCTCTGCCTGGAGATGGACGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGT  460

Db     270  CTCCTGCAGAACAACTATGGACTTGCCAGTTCAAAAGTTTCAAAAGTTTCCTGAGTTCTGAATTCAGT  329

Qy     461  GAGGAGAACAATTGAATTTGGGTGCCCTGTGAGACATCGCCCAGCCCACT  520
```

```
Db     609  GAATTCTACCAGGACTTG  626
                 ||    ||||    ||||
```

RESULT 8
US-08-748-483-2
; Sequence 2, Application US/08748483
; Patent No. 5955314
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: A NOVEL REGULATOR OF CELL SIGNALLING
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 Porter Drive
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: US
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/748,483
;   FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Billings, Lucy J.
;   REGISTRATION NUMBER: 36,749
;   REFERENCE/DOCKET NUMBER: PF-0157 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-855-0555
;   TELEFAX: 415-845-4166
;   TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 744 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   LIBRARY: Consensus
;   CLONE: 57362
US-08-748-483-2

Query Match   12.2%;   Score 142; DB 2; Length 744;
Best Local Similarity 59.8%; Pred. No. 1e-26;
Matches 238; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

Qy   341 GCCAAAGAAACAAGAGTCTCTCCTGAAGAAGCAGTGAAATGGCTGAATCATTGACAAA 400
                 |||||| ||||||||||||| |||||||
Db   210 GCCAAGACCCAGAAAAACCTCGCTGGACGAGGCCCTGCAGTGGCGTGGGTGATTCCCTGGACAAA 269

Qy   401 TTGCTCTCTCATAGAGATGGAGATGGATGCTTTTAAAACTGAATTCAGT 460
                 ||||| ||| ||| |||||||| || ||||| ||||| |||||
Db   270 CTCCTGCAGAACAACTATGGGACTTGCCAGTTCAAAAGTTCTGAATTCAGT 329

Qy   461 GAGGAGAACAATTGAATTTGGGTGCCGCTGCGTGAGACTTGCAAGGACACCTCAA 520
                 |||||||| ||| || ||||||||||||||||||||||||||||||
Db   330 GAGGAAAACCTTGAGTTCTGAATTGCCTGTCGTGAGGATTACAAGAAGATCAAGTCCCCTGCC 389

Qy   521 CAAATCATCCTAAAAGCAAAGGCAATCTATGAGAATTCTTCAGAATGATGCCCCAAA 580
                 ||| ||| ||||||||||||||| ||| || || || |||| ||| |||||
Db   390 AAGATGGCTGAGAAGGCAAAGCAAAATTATGAGAATTCATTCAAACGGAGGCTCTAAA 449

Qy   581 GAGGTTAACATTGATTTCATCTAAAGAAGAAGTAATTGCTAAGAGCATGCTCCAGCCCACT 640
                 || || |||| || || || || |
```

```
Db   140 AGTGCAAAGATGAAACATCGGCTAGGTTTCCTGCTGCAAAAATCGATTCCTGTGAA 199
Qy   311 GAGA---CTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAACAAGAGTCTCTCCTGAA 367
Db   200 CACAATTCTTCCCACACAAGAGAGGACAAAGTGGTTATTTGCCAGAGTGAGCCAAGAG 259
Qy   368 GAAGCAGTGAAATGGGCTGATCATTTGACAAATTCTCTCTCATAGAGATGGAGTGGAT 427
Db   260 GAAGTCAAGAAATGGGTGGAATCACTCGAAAACCTGATTAGTCATCATGATGTGGGCTCGCA 319
Qy   428 GCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTGGGTGCCC 487
Db   320 GCTTTCAAAGTTCTTGAGTCTGAATATAGTGAGGAGAATATTGACTTCTGGATCAGC 379
Qy   488 TGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAAGGCAATC 547
Db   380 TGTGAAGAGTACAAGAAAATCAAATCACCATCTAAACTAAGTCCCAAGTTTAACATTGAT 439
Qy   548 TATGAGAAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTTCATACTAAA 607
Db   440 TATAATGAATTCATCTCAGTCCAGGCAACCAAAGAGGTGAACCTGGATTTCTGCACCAGG 499
Qy   608 GAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCCACAGTTTGTGATACGGCACAAAGC 667
Db   500 GAAGAGACAAGCCGGAACATGCTAGAGCCTACAATAACCTGCTTTGATGAGGCCCAGAAG 559
Qy   668 AGAGTGTACCAGCTCATGGAACATGACAGTTATAAATCTGAAATCTGAGACCTAC 727
Db   560 AAGATTTTCAACCTGATGGAGAGAGGATTCCTACCGCGCTTCCTCAGTCTCTGATTCTAT 619
Qy   728 TTACATTTGAT 738
Db   620 CTTGATTTGGT 630
```

RESULT 11
US-09-566-921-111
; Sequence 111, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
;- APPLICANT: Loring, Jeanne F.
;- APPLICANT: Tingley, Debora W.
;- APPLICANT: Edwards, Carla M.
;- TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
;- FILE REFERENCE: PA-0024 US
;- CURRENT APPLICATION NUMBER: US/09/566,921
;- CURRENT FILING DATE: 2000-05-05
;- NUMBER OF SEQ ID NOS: 138
;- SOFTWARE: PERL Program
; SEQ ID NO 111
;- LENGTH: 3208
;- TYPE: DNA
;- ORGANISM: Homo sapiens
;- FEATURE:
;- NAME/KEY: misc feature
;- OTHER INFORMATION: Incyte ID No. 6682888 900146.4
US-09-566-921-111

Query Match   11.7%;   Score 136; DB 4; Length 3208;
Best Local Similarity 60.6%;   Pred. No. 5.8e-25;
Matches 223;   Conservative 0;   Mismatches 145;   Indels 0;   Gaps 0;

```
Qy   371 GCAGTGAAATGGGCTGAATCATTTGACAAATTGCTCTCATAGAGATGGAGTGGATGCT 430
Db   321 GCCCTGCAGTGCGTGATTCCCTGAGCAAATTCCTGCAGGAACAACTATGGACTTGCCAGT 380
Qy   431 TTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTGGGTCGCCTGT 490
Db   381 TTCAAAGTTCTTCCTGAGTCTGAATTCAGTCAGTGAGGAGAAAACCTTGAGTTCTGGTT 440
Qy   491 GAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAAGGCAATCTAT 550
```

```
Db   330 GAGGAAAACCTTGAGTTCTGGATTGCCTGTGAGGATTACAAGAGATCAAGTCCCCTGCC 389
Qy   521 CAAATCATCCTAAAAGCAAAGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAA 580
Db   390 AAGATGGCTGAGGAGGCAAAGCAAATTTATGAGAATTCATTCAAACGGAGGCTCCTAAA 449
Qy   581 GAGGTTAACATTGATTTTCATCTAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACT 640
Db   450 GAGGTGAATATTGACCACTTCACTCAAGGACATCACAATGAAGAACCTCGTGGAACTTCC 509
Qy   641 CTCCACAGTTTGTATACGGCACAAGAGCAGAGGTACCAGCTCCATGGACACATGACAGTTAT 700
Db   510 CTGGACAGCTTTGACATGCCCAGAAAAGAAATCCATGCCCTGATGGGAATAATTCTG 569
Qy   701 AAACGCTTTTGAAAATTGAGACCTACTTACATTTGAT 738
Db   570 CCTCGCTTTGTGCGCTCTGAGTTTTATCAGGAGTTAAT 607
```

RESULT 10
US-09-016-434-1091
; Sequence 1091, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
;- APPLICANT: Janice Au-Young
;- APPLICANT: Jeffrey J. Seilhamer
;- TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
;- TITLE OF INVENTION: PATHWAY GENE EXPRESSION
;- NUMBER OF SEQUENCES: 1490
;- CORRESPONDENCE ADDRESS:
;- ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;- STREET: 3174 PORTER DRIVE
;- CITY: PALO ALTO
;- STATE: CALIFORNIA
;- COUNTRY: USA
;- ZIP: 94304
;- COMPUTER READABLE FORM:
;- MEDIUM TYPE: Floppy disk
;- COMPUTER: IBM PC compatible
;- OPERATING SYSTEM: PC-DOS/MS-DOS
;- SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;- CURRENT APPLICATION DATA:
;- APPLICATION NUMBER: US/09/016,434
;- FILING DATE: HEREWITH
;- CLASSIFICATION:
;- PRIOR APPLICATION DATA:
;- APPLICATION NUMBER:
;- FILING DATE:
;- CLASSIFICATION:
;- ATTORNEY/AGENT INFORMATION:
;- NAME: Zeller, Karen J.
;- REGISTRATION NUMBER: 37,071
;- REFERENCE/DOCKET NUMBER: PA-0002 US
;- TELECOMMUNICATION INFORMATION:
;- TELEPHONE: (650) 855-0555
;- TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1091:
;- SEQUENCE CHARACTERISTICS:
;- LENGTH: 800 base pairs
;- TYPE: nucleic acid
;- STRANDEDNESS: single
;- TOPOLOGY: linear
;- IMMEDIATE SOURCE:
;- LIBRARY: GENBANK
;- CLONE: g1216372
US-09-016-434-1091

Query Match   11.9%;   Score 139; DB 4; Length 800;
Best Local Similarity 56.6%;   Pred. No. 6e-26;
Matches 278;   Conservative 0;   Mismatches 210;   Indels 3;   Gaps 1;

```
Qy   251 AGAGCGAAAGAAAAAAGGAATAACTAAGTCTTCCTCCTCACACGAGAGGCCTGACTTCCATGGA 310
```

```
Db  1318 AGAGCAAAACCTAGCACGAGAATCCCCTGCCAAGATGAAGAACAAGCTGGGGATCTTCAGACGCGGCGGAATG 1377
Qy   300 ACTTCCATGGAGAGACTCAAGCCAGCCAGTAGATCTGCCCTCTTGGCCAAGAAACAAGAGTCT 359
         |||||||||||||||
Db  1378 AGTCCCTGGAGCCCTCCCGCGGGCAAGCAGACAAAATGATGAAGTCATTCAGCCCA 1437
Qy   360 CTCCTGAAGAAGCAGTGAAATGGCTGAATCATTTGACAAATTGCTCTCATAGAGATG 419
         |||||||
Db  1438 CCTCAGAGGAAGCCCTCAGTGGGCGAGTCCTTGGAGAAGCTGCTGGTTCACAAATACG 1497
Qy   420 GAGTGGATGCTGTTTACCAGATTTCTAAAACTGAATTCAGTGAGAGAACATTGAATTTT 479
         |||||||||
Db  1498 GGTTAGCAGTGTTCCAAGCCTTCCTTCGCACTGAGTTCAGTGAGTCGGAGTTCT 1557
Qy   480 GGGTCGCCCTGTGAGCACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAA 539
         |||||||||
Db  1558 GGTTGGCTTGTGGAGGACTTCAAGAAGGTCAAGTCAAGATGGCATCCAAGGCCA 1617
Qy   540 AGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTTC 599
         |||||||||||
Db  1618 AGAGATCTTTGCTGAATACATCCGCATCCAGGCATCGCCCACTCTCCACAGTTTTGATACGG 1677
Qy   600 ATACTAAAGAGAGTAATTGCTAAGAGCATCGAAGGAGTCACCTGGACTCCT 659
         |||||||||
Db  1678 ACACGCGGGGAGCACACCAAGGACAACCTGCAGAGCGTCACGCGGGGCGTCGTTCGACCTGG 1737
Qy   660 CACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAAATCTG 719
         |||||||||
Db  1738 CACAGAGAGCGCATCTTCGGGCTCATGCATGGAAAAGGACTCGTACCCTCGCTTTCTCCGTTCTG 1797
Qy   720 AGACCCACTTACATTTGAT 738
         |||||||
Db  1798 ACCTTTACCTGGACCTTAT 1816
```

RESULT 13

```
Qy   441 GAGGATTACAAGAAGATCAAGTCCCCTGCCAAGATGGCTGAGAAGGCAAGCAAATTTAT 500
         |   |||  || ||
Qy   551 GAGAAAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTTCATACTAAAGAA 610
         |  |||||||| ||
Db   501 GAAGAATTCATTCAAACGGAGCTCCTTAAAGAGGTGAATATTGACCACTTCACTAAGGAC 560
Qy   611 GTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGATACGGCACAAAGCCAGA 670
         ||   |  ||
Db   561 ATCACAATGAAGAAGACCTGTGGAACCTTCCCTGGACCAGCTTTGACATGGCCCAGAAAAGA 620
Qy   671 GTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAAATCTGAGACCTACTTA 730
         |||  |  ||
Db   621 ATCCATGCCCTGATGGAAAAGGATTCTCTGCCCTGCTTTTGTCGCGTCTGAGTTTTATCAG 680
Qy   731 CATTTGAT 738
         |||||
Db   681 GAGTTAAT 688
```

RESULT 12

```
Query Match           11.4%;  Score 132.6;  DB 4;  Length 2638;
Best Local Similarity 54.1%;  Pred. No. 3.9e-24;
Matches 270;  Conservative 0;  Mismatches 229;  Indels 0;  Gaps 0;

Qy   240 AGGCCAAAATCAGACGAGCGAAAGAAAGGAATAGACTAAGTCTTCTCCTACAGAGGCCTG 299
         ||  ||||||| ||||  |||| ||||
```

SEQUENCE CHARACTERISTICS:
; LENGTH: 1398 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g297704
US-09-016-434-1290

Query Match   11.2%; Score 129.8; DB 4; Length 1398;
Best Local Similarity   60.2%; Pred. No. 1.6e-23;
Matches 234; Conservative   0; Mismatches 152; Indels   3; Gaps   1;

```
Qy   355 AGTCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAG 414
             |||||||||||||||||||||    ||| |||||||||| ||  ||||||  ||
Db   278 ACTTTCTGCTGCTGAAGTAATGCAATGGTCTCAATCTCTGGAAAAACTTCTTGCCAACCA 337

Qy   415 AGATGGAGTGGATGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGA 474
             || ||| ||||||||  |||||  ||||||  ||||||||||||||||||||||||
Db   338 AACTGGTCAAAATGTCTTTGGAAGTTCCTAAAGTCTGAATTCAGTGAGGAGAATATTGA 397

Qy   475 ATTTTGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAA 534
             ||||||||||||| |||||||| |||||||   |||||||  ||||||||||||
Db   398 GTTCTGGCTGGCTTGTGAAGACTATAAGAAA---ACAGAGTCTGATCTTTTGCCCTGTAA 454

Qy   535 AGCAAAGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGA 594
             |||||||||| ||||||||||||||||  |||||||  ||||||||||  |||||||
Db   455 AGCAGAAGAGATATATAAAGCATTTGTGTCATTCAGATGCTGCTAAACAAATCAATATTGA 514

Qy   595 TTTTCATACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGA 654
             ||||  ||| ||||     |||||||| |||||||||||||||||||||||||||||
Db   515 CTTCCGCACTCGAGAATCTACAGCCAAGAAGAATTAAAGCACCAACCCCACGTGTTTTGA 574

Qy   655 TACGGCACAAAGCAGAGTGTACCAGCTCATGGGACAATGACAGTTATAAACGCTTTTTGAA 714
             |||||||||||  || ||| ||||||||||  ||||  ||| ||||||  |||| |||
Db   575 TGAAGCACAAAAGTCATATATACTCTTATGGGAAAAGGACTCTTATCCCAGGTTCCTCAA 634

Qy   715 ATCTGAGACCTACTTACATTTGATAGAAG 743
             ||||||||  |||| |||||||||||
Db   635 ATCAGATATTTACTTAAATCTTCTAAATG 663
```

RESULT 14
US-09-702-705-322
; Sequence 322, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 322
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-322

Query Match   11.2%; Score 129.8; DB 4; Length 1398;
Best Local Similarity   60.2%; Pred. No. 1.6e-23;
Matches 234; Conservative   0; Mismatches 152; Indels   3; Gaps   1;

RESULT 15
US-09-736-457-322
; Sequence 322, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 322
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-322

Query Match   11.2%; Score 129.8; DB 4; Length 1398;
Best Local Similarity   60.2%; Pred. No. 1.6e-23;
Matches 234; Conservative   0; Mismatches 152; Indels   3; Gaps   1;

```
Qy   355 AGTCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAG 414
             |||||||||||||||||||||    ||| |||||||||| ||  ||||||  ||
Db   278 ACTTTCTGCTGCTGAAGTAATGCAATGGTCTCAATCTCTGGAAAAACTTCTTGCCAACCA 337

Qy   415 AGATGGAGTGGATGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGA 474
             || ||| ||||||||  |||||  ||||||  ||||||||||||||||||||||||
Db   338 AACTGGTCAAAATGTCTTTGGAAGTTCCTAAAGTCTGAATTCAGTGAGGAGAATATTGA 397

Qy   475 ATTTTGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAA 534
             ||||||||||||| |||||||| |||||||   |||||||  ||||||||||||
Db   398 GTTCTGGCTGGCTTGTGAAGACTATAAGAAA---ACAGAGTCTGATCTTTTGCCCTGTAA 454
```

```
Qy    535  AGCAAAGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGA  594
                ||||-|-|||-|-|||-| |||| -|||| --|||| -| ||-|||||
Db    455  AGCAGAAGAGATATATAAAGCATTTGTGCATTCAGATGCTGCTAAACAAATCAATATTGA  514

Qy    595  TTTTCATACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGA  654
                ||||-|||| |-|| |||||-| --|-||-| ||||--|||||
Db    515  CTTCCGCACTCGAGAATCTACAGCCAAGAAGATTAAAGCACCACCCCCACGTGTTTTGA  574

Qy    655  TACGGCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAA  714
                ||||||-||||| -|||| -|-||||-|||| -|-|||-| ||||
Db    575  TGAAGCACAAAAGTCATATATACTCTTATGGAAAAGGACTCTTATCCCAGGTTCCTCAA  634

Qy    715  ATCTGAGACCTACTTACATTTGATAGAAG  743
                |||-||-|||||||-|-|| |||| ||
Db    635  ATCAGATATTTACTTAAATCTTCTAAATG  663
```

Search completed: August 20, 2004, 12:36:26
Job time : 94.2662 secs

```
            GenCore version 5.1.6
      Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:     August 20, 2004, 09:59:05 ; Search time 543.613 Seconds
                                         (without alignments)
                                    10518.000 Million cell updates/sec

Title:           US-09-894-749-3
Perfect score:   1164
Sequence:        1 ttttgtaagaaaatctga........gtactagtggatccgagctc 1164

Scoring table:   IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched:        3228839 seqs, 2456066551 residues

Total number of hits satisfying chosen parameters:    6457678

Minimum DB seq length:  0
Maximum DB seq length:  2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :    Published Applications NA:*
        1:   /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
        2:   /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
        3:   /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
        4:   /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
        5:   /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
        6:   /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
        7:   /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
        8:   /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
        9:   /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
       10:   /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
       11:   /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
       12:   /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
       13:   /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
       14:   /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
       15:   /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
       16:   /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
       17:   /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
       18:   /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
       19:   /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
            /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1164	100.0	1164	9	US-09-894-749-3	Sequence 3, Appli
2	602.6	51.8	1884	17	US-10-115-635-133	Sequence 133, App
3	599.6	51.5	1840	13	US-10-258-371B-18	Sequence 18, Appli
4	599.6	51.5	2144	13	US-10-258-371B-19	Sequence 19, Appli
5	596	51.2	2217	9	US-09-894-749-1	Sequence 1, Appli
6	560.6	48.2	848	9	US-09-867-550-847	Sequence 847, App
7	305.4	26.2	736	9	US-09-833-381-1010	Sequence 1010, Ap
8	192.2	16.5	1486	13	US-10-258-371B-6	Sequence 6, Appli
9	177.6	15.3	241	13	US-10-258-371B-11	Sequence 11, Appl
10	152.4	13.1	1345	13	US-10-342-887-839	Sequence 839, App
11	152.4	13.1	1345	13	US-10-172-118-839	Sequence 839, App
12	152.4	13.1	1345	16	US-10-305-720-1278	Sequence 1278, Ap
13	152.4	13.1	1345	17	US-10-641-643-1243	Sequence 1243, Ap
14	152.4	13.1	1364	9	US-09-925-300-567	Sequence 567, App
15	152.4	13.1	1381	10	US-09-971-429B-32	Sequence 32, Appl
16	152.4	13.1	1382	10	US-09-971-392-13	Sequence 13, Appl
17	152.4	13.1	1629	12	US-10-152-319A-1995	Sequence 1995, Ap
18	152.4	13.1	1629	16	US-10-191-803-244	Sequence 244, App
19	143.6	12.3	540	15	US-10-275-555-1	Sequence 1, Appli
20	143.6	12.3	939	13	US-10-302-172-353	Sequence 353, App
21	142	12.2	543	15	US-10-398-953-2	Sequence 2, Appli
22	142	12.2	744	9	US-09-206-639-2	Sequence 2, Appli
23	142	12.2	817	15	US-10-338-953-6	Sequence 6, Appli
24	142	12.2	1923	16	US-10-411-224-55	Sequence 55, Appl
25	142	12.2	1923	13	US-10-047-021-55	Sequence 55, Appl
26	142	12.2	2076	13	US-10-240-425-345	Sequence 345, App
27	142	12.2	2190	17	US-10-641-643-795	Sequence 795, App
28	142	12.2	5809	15	US-10-133-013-103	Sequence 103, App
29	140	12.0	1283	10	US-09-814-353-19578	Sequence 19578, A
30	139	11.9	800	16	US-10-305-720-1091	Sequence 1091, Ap
31	139	11.9	1531	15	US-10-198-846-9863	Sequence 9863, Ap
32	139	11.9	1869	15	US-10-175-523-144	Sequence 144, App
33	139	11.9	2752	15	US-10-198-846-13843	Sequence 13843, A
34	139	11.9	2934	10	US-09-939-209A-1	Sequence 1, Appli
35	139	11.9	3165	14	US-10-044-090-113	Sequence 113, App
36	139	11.9	3167	15	US-09-918-624B-6	Sequence 6, Appli
37	133.8	11.5	420	9	US-09-960-352-2229	Sequence 2229, Ap
38	132.6	11.4	2623	16	US-10-108-260A-1527	Sequence 1527, Ap
39	132.6	11.4	2638	10	US-09-954-456-1955	Sequence 1955, Ap
40	132.6	11.4	2638	10	US-09-873-367C-1040	Sequence 1040, Ap
41	132.6	11.4	2638	13	US-10-240-425-1426	Sequence 1426, Ap
42	132.6	11.4	2638	16	US-10-428-487-13	Sequence 13, Appl
43	132.6	11.4	2638	16	US-10-305-720-1090	Sequence 1090, App
44	129.8	11.2	1329	15	US-10-247-671-107	Sequence 107, App
45	129.8	11.2	1380	9	US-09-925-301-450	Sequence 450, App

ALIGNMENTS

```
RESULT 1
US-09-894-749-3
; Sequence 3, Application US/09894749
; Patent No. US20020081683A1
; GENERAL INFORMATION:
;  APPLICANT: Hodge, Martin R.
;  APPLICANT: Yowe, David
;  TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
;  FILE REFERENCE: 5800-19, 035800/17480
;  CURRENT APPLICATION NUMBER: US/09/894,749
;  CURRENT FILING DATE: 2001-06-27
;  PRIOR APPLICATION NUMBER: 09/244,314
;  PRIOR FILING DATE: 1999-02-04
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;  LENGTH: 1164
;  TYPE: DNA
;  ORGANISM: Mus sp.
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (134)..(841)
US-09-894-749-3

  Query Match         100.0%;  Score 1164;  DB 9;  Length 1164;
  Best Local Similarity   100.0%;  Pred. No. 9e-242;
  Matches 1164;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 TTTTTGTAAGAAAAATCTGAGGAAGATTCGGGATAGCGCTTTATTCAGGATGTTTTCCT 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTGTAAGAAAAATCTGAGGAAGATTCGGGATAGCGCTTTATTCAGGATGTTTTCCT 60

Qy     61 ATGAAATAGCATTCATCTGTGGGAGAGAAGGAGGACTAAGGAAATCTGACATCTGTGGTC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATGAAATAGCATTCATCTGTGGGAGAGAAGGAGGACTAAGGAAATCTGACATCTGTGGTC 120

Qy    121 ACTGGGACAGAATATGGATATGTCACTGGTTTCTCTCCTCAATTAAATATGTGTGAATC 180
```

US-10-115-635-133
; Sequence 133, Application US/1011635
; Publication No. US20040137434A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y., Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 797CON
; CURRENT APPLICATION NUMBER: US/10/115,635
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 09/714,936
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 362
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 133
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (157)..(864)
US-10-115-635-133

Query Match   51.8%; Score 602.6; DB 17; Length 1884;
Best Local Similarity 75.3%; Pred. No. 3.7e-120;
Matches 817;   Conservative   0; Mismatches 234; Indels   34; Gaps   4;

Qy   1 TTTTTGTAAGAAAATCTGAGGAAGATTCGGGATAGCGCTTTATTC-AGGATGTTTTCC 59
            || ||| ||| || |||||||||| || |||||||| |||| ||||| |||| ||||
Db   25 TTCTTTTTGTAAACATTACTGTAAGAGTTGTGATAACTTTTATTCTACTATGTATATG 84

Qy   60 TATGAAATAGCATTCATCTGTGGGAGAGAGAGGACTAAGGAAATCTGACATCTGTTGGT 119
            |||||||||| ||||| ||||||||||||||||||| |||||||||||||||| |||||
Db   85 TATGGAATAGTATTAATAAATGAACTAGGAGGATGTAATAAATTAGACATCTCTTCAT 144

Qy   120 CACTGGGACAGAATATGGATATGTCACTGGTTTTCTTCTCTCAATTAAATATGTGTGAAT 179
            |||||||| |||||||||||| ||||| ||| ||||| |||| ||| |||||||||
Db   145 T--TTAGAGGAAGATGGAAACAACATTGCTTTCTTCTTTTCAAATAAATAATGTGTGAAT 202

Qy   180 CAAAAGAGAAAACTTTTTCAAACTAATGCAGGTCAGGGAAAAGAAGAAGACAAGCATCG 239
            |||||||||||| ||||||||||| ||||| ||||||| ||||||||||||| ||||
Db   203 CAAAAGAAAAAACTTTTTCAAGTTAATACATGGTTAATACATAGGTTCAGAGAAACAAGCAAAG 262

Qy   240 AGGCCAAAATCAGAGCGGCAAGAAGAAAGGAATAAGACAGAGATAGACTAGTCTTCCTACGAGAGCCTG 299
            ||||||||||||||||| |||||||| ||||| ||||||| |||| |||||||||| ||||||| |||||
Db   263 AAGCCAAAATCAGAGCGTCAAGGAATGAGCAGTAAGAAAGAAAGAAAGAAGAAAAGTAAGACTAAGTCCTGTGCAGAAACCCTG 322

Qy   300 ACTTCCATGGAAGAGAGACTCAAGCCAGTGAGGTCGAAATGGGCTGAATCATTGACAAATTGCTTCTCTCATAGAGATG 359
            |||| ||||||||||||| |||||||||||| ||||||||||||||| |||||||||||| |||||||||| ||||||
Db   323 AGTTTCATGGAAGAGAGAGACTCGGGCACTTGCCAGTAGATCTGGGCACTTGCCAGACTCGGGCACTTGCCAGTGGCTGAATCATTTGACAAACTGCTTTGCCAGAGAAATG 382

Qy   360 CTCTGAAGAAGCAGTCAGTGAAATGGGCTGAATCATTGACAAATTGGTCTTCCAGAATTGCTAATA 419
            ||| |||||||| ||| ||||||| |||||| ||||||| ||| |||||||| ||||||||||||||
Db   383 CCCCTGAAGAGCGCAGTTCAGTGAAATGGGGTGAATCGTTGGGTGATTGAGGTCGTTTCCCATGAGAATG 442

Qy   420 GAGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAACATTGAATTTT 479
            ||| |||||||| || ||||||||| |||||||||| ||||| ||||||||||||||||
Db   443 GACTAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAAGAAAATTGAATTTT 502

Qy   480 GGGTCGCCCTGTGAAGACTTCAAGAAATGCAAGACCTCAACAAATCATCCTAAAAGCAA 539
            || || | |||||||| ||||||||||||||||| |||||||||||||| |||||||||
Db   503 GGATAGCGCTGTGAAGAGTTCAAGAAAAGCAAGACGGACCTCAACAAATTCACCTTAAAGCAA 562

RESULT 2

```
Qy    63  GAAATAGCATTCATCTGTGGGAGAGAAGGACTAAGGAAATCTGACATCTGTTGGTCAC  122
              |||||||||||||| |||| |||| ||| |||||| |  | |||||||| |||| ||
Db    94  GGAATAGTATTAATAAATGACTAGGAAGGAATGTAATAAAATTAGACATCTTCATT--  151

Qy   123  TGGGACAGAAATGATGATGTCACTGGTTTCTCTCAATTAAATAATGTGATCAA  182
              |||||| ||||| |||||||||| ||||||||||||||||||||| |||||||
Db   152  TTAGAGAGAAATGAAACACATTGCTTTTCTTCTTCTCAATAAATAATGTGATCAA  211

Qy   183  AAGAGAAAACTTTTTCAAACTAATGCCATGGTCAGGGAAGAAGAAGAACCATCGAGG  242
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   212  AAGAAAAAACTTTTTCAAGTTAATACATGTTCAGGGAAAGAAGAAGAACAAGCAAGAAG  271

Qy   243  CCAAAATCAGAGCGAAAGAAAAAAGGAATAGACTAAGTCTTCTCCCACAGAGGCCTGACT  302
              ||||||||||||||||||||||||||||||| |||||||||||| ||||||||||||
Db   272  CCAAAATCAGAGCAGCTAGGAAGAAAAATAGAGAATAGAGTCTTCTTGTGTGCAGAAACCTGAGT  331

Qy   303  TCCATGGAGACTCAAGCCAGTAGAACTCGCCCTTTCGCCAAAGAAACAAGAGTCTCTC  362
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   332  TTCATGAAGACACACCCGCTCCAGTAGATCGGCCACTCGGCCAAAGAAACAAGAGTCTCCC  391

Qy   363  CTGAAGAGCAGTGAAATGGGCTGAATCATTGACAAATTGCTCTCTCATGAGATGGGAG  422
              ||||| |||||||||||||| ||||||| |||||||||||||||| ||||||||||||
Db   392  CTGAAGAGGCAGTGAAATGGGGTGAATCATTTGACAAAC TGCTTTCCCATAGAGATGGAC  451

Qy   423  TGGATGCTTTACCAGATTCTTAAAAATAGCTTGTGTGATAAGAAGAGATGAATTTCATTTT---  482
              |||||||||||||||||||||||||| |||| ||||||||||| |||||||||||||||||
Db   452  TAGAGGCTTTTACCAGATTCTTAAAAATGAATTCAGTGAAGAAAAATATTGAATTTTTGGA  511

Qy   483  TCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAACATCCTAAAAGCAAAGG  542
              |||||||||||||||||||||||||||||||||| ||||||||||||| |||||||||
Db   512  TAGCCTGTGAAGATTCAAGAAATGCAAGGAGAGACCTCAACAAATCTTAGGAGCAAAG  571

Qy   543  CAATCTATGAGAAATCATTCAGAAATGATGCCCCAAAGAGTTAACATTGATTTTCATA  602
              |||||||||||||| ||||||||| ||||| |||||||||||||||||||||||||||
Db   572  CAATATATGAGAAATTTATACAGACTGATGCGCCCAAAAGAGTTAACCTCGATTTTCACA  631

Qy   603  CTAAAAGAAGTAATTGCTAAGAGCATCGCGCCAGCCCACTCCTCCACAGTTTGATACGGCAC  662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   632  CAAAAGAAGTCATTACAAAACAGCATCACTCACCCCTCACCCTCCACAGTTTGATGCGTGCAC  691

Qy   663  AAAGCGAGAGTGACCAGCTCATGGAACATGACATGACAGCAGTTATAAACGCTTTTGAAATCTGAGA  722
              ||||||| |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   692  AAAGCAGAGTGTATCAGCTCATGGAACAAGACCAGTCATACACGTTTTCTGAAATCTGACA  751

Qy   723  CCTACTTACATTTGATAGAAGGAGAACCTCAGAGACCTCAGAGACGATCAC  782
              |||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   752  TCTATTTAGACTTGATGGAAGGAAGACCTCAGAGACGACCACCAAATCTTAGGAGACGATCAC  811

Qy   783  GATCATTTACTTACAATGATTCCAAGATGTAAAGTCAGATGTTGCCATTGGTTATGAG  842
              ||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||
Db   812  GCTCATTTACCTGCCATGGCCAATGAATTCCAAGATGTACAAATCAGATGTTGCCATTGGTTATAAA  871

Qy   843  TAAAAGTCATTGTCTTCTTTTGTGATAGTGTATGTGTAAAATATACTAATAATACT  902
              ||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   872  GAAAATTGATTTTGCTCATTTTATGACAAACTTATACATC---  912

Qy   903  AATGTGTACTTCTAAAATATAGCTTGTGTGATAGAAGAAGATGATTTCATTTT------TA  956
              ||||||||||| |||||||||||||||||||||| |||
Db   913  ------TGCTTCTCAACATATCGCATGTTATGTTAAGATTTGGTCCCATCCTTTAAACTG  966

Qy   957  AAATACACCATGCAAATACCAAATACATATTAAAATGTAAAAATCTATCTGATATTCTAATAAAAT  1016
              ||||   |||||||||| ||||||||||||||||||||||||||||| |||
Db   967  AAATAATGTCATGTGAAATAATTTTTAAAAAATGTAAAAACAAAACTTTCTGCTAACAAAATA  1026

Qy  1017  CATCATCTATCTTCCGAAATATTTTATGAAAATCTATCTGATATTCTAATAAAAT  1076
              |||    ||  |  ||| |||   |||| |||| | |||| ||| | ||
Db  1027  CATACAGTATCTGCCAGTATATTCTGCTAAAACCTTCTGATTCATTTGATGTCATTCATTTATAA  1086

Qy  1074  AATTC  1078
              |||||
Db  1078  TAATC  1082

Qy  1077  TC  1078
Db  1087  TC  1088
```

```
Qy     4  TTGTAAGAAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTC-AGGATGTTTTCCTAT  62
             ||  ||||  |||||     | ||||  |||  ||| ||| |||  |||||||| ||
Db    34  TTTTTGTAAACATTACTGTAAGAGTTGTGATAACTTTTTATTCTACTATGTATATGTAT  93
```

RESULT 4
US-10-258-371B-19
; Sequence 19, Application US/10258371B
; Publication No. US20040067903A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS-GAGNON, Alison
; APPLICANT: MURRAY, David L
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING,
; TITLE OF INVENTION: RGS18, AND USES THEREOF
; FILE REFERENCE: A3656 US PCT
; CURRENT APPLICATION NUMBER: US/10/258,371B
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: GB001883.334
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US60/200,786
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 2144
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-258-371B-19

Query Match   51.5%;   Score 599.6;   DB 13;   Length 2144;
Best Local Similarity   75.2%;   Pred. No. 1.8e-119;
Matches 814;   Conservative   0;   Mismatches 234;   Indels   34;   Gaps   4;

```
Qy   4   TTGTAAGAAAATCTGAGGAAAGATTCGGGATGAGCGGCTTTATTC-AGGATGTTTCCTAT   62
         |||||||  ||   ||  |||||  |  ||    ||  |||||||| ||| ||| ||| |
Db  34   TTTTTTGTAAACATTACTGTAAGAGTTGTGATTAACTTTTTATTCTTACTATGTATATGTAT   93

Qy  63   GAAATAGCATTCATCTGTGGGAGAGAGGAAGACTAAGGAAATCTGACATCTGTTGGTCAC  122
         | ||||| ||| || |  |||  ||| ||||  || | |||||||||  | ||  || ||
Db  94   GGAATAGTATTAATAAATGAACTAGGGAAGGATGTAATGAAAATCTGATAAAATAGTGAATCAA--  151

Qy 123   TGGGACAGAATATGGATAATGTCACTGGTTTCTTCTCTCAATTAAATAATGTGTGAATCAA  182
         | ||| || ||| ||| |   |   ||||||||| ||||||| ||||||||||||||||||
Db 152   TTAGAGAAGATGGAAACAACATTGCTTTTCTTTTTTCTCAAATAAATAATGTGTGAATCAA  211

Qy 183   AAGGAGAAAACTTTTTCAAATGCATGGGTCAGGGAGTCAGGGTCAAGCATCGAGG  242
         |||  |||||||||||||||  | || || ||||| |||||  |  ||| | | |
Db 212   AAGAAAAAAACTTTTTTCAAGTTAATACATGGTTCAGGATCAGGGAAAGAAGAAGCAACAGAAG  271

Qy 243   CCAAAATCAGACGCGAAAGAAGAAAAAGAATAGACTAAGTCTTCTCCTACAGAGGCCTGACT  302
         |||||||||||  | ||||| |||||||| ||||||| || |||  ||  ||||| | ||||
Db 272   CCAAAATCAGAGCTAAGGAAGCTAAGGAAAAAGAAAATAGACTAAGCTTCTTGTGTCAGAAACCTGAGT  331

Qy 303   TCCATGCTGAGAGACTCAAGCCAGTAGACTCGCCCTTCTTGGCCAAAGAAACAGAGTCTCTC  362
         | ||| | | ||  |  |||||||||| || |  || | ||||||||||||||| ||||| |
Db 332   TTCATGACAAACCCGCTCCAGTAGATCGGGCACTTGGGCCAAAGAAACAAGAGTCTCCC  391

Qy 363   CTGAAGAGCAGTGAAATGGGCTGAATCATTGACAAAATTGCTCTCTCATAGAGATGGAG  422
         |||||||| |||||||||||  |||||||||||||| ||||| | ||||||||||||| |
Db 392   CTGAAGAGAGGCAGTGAAATGGGGGTGAATCATTGACAAAACTGCTTTCCCATAGAGATGGAC  451

Qy 423   TGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTGGG  482
         | || || ||||||||||||||||||||||||||||||||| | |   |||||||||| |
Db 452   TAGAGGCCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAAGAAATATTGAATTTTGGA  511

Qy 483   TCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAAGG  542
         | ||||| ||||| |||||||||||||||||||||||||||| | ||| ||| |||| |
Db 512   TAGCCTGGTGAAGATTTCAAGAAATGCAAGGAACCTCAACAAATTCACCTTAAAGCAAAAG  571

Qy 543   CAATCTCATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTTCATA  602
         |||| | |||||||||| || |||| |||||||||||| |||||||||  | |||||||
Db 572   CAATATATGAGAAATTTATACAGACTGATGCCCCCAAAAGAGGTTAACCTCGATTTTCACA  631

Qy 603   CTAAAGAAGTAATTGCTAAGAGCCATCGCCCCAGCCACCTCCCACAGTTTGTATACCGCGAC  662
         | |||||||| ||| | |||| ||||| |  ||||| |  |||||||||||||| |  | ||
Db 632   CAAAAGAAGTCATTACAACAGCACTCACTCAACCTACCCTCCCACAGTTTTGATGCTGCAC  691
```

```
Qy 663   AAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTGAAATCTGAGA  722
         ||||||||||||| |||||||||||||| | ||||||||| || |||| ||||||||| |
Db 692   AAAGCAGAGTGTATCAGCTCATGGAACACAAGACAGTTATCACGTTTTCTGAAATCTGACA  751

Qy 723   CCTACTTACATTGTATAGAAGGAGAACCTCAGAGACCAACAAACCTTAGGAGACGATCAC  782
           | | | |  ||||||| |||||||||||||||||||||||| |||||||||||||||||
Db 752   TCTATTTAGACTTGATGGAGGAAGGAGACCTCAGAGACCAACAAATCTTAGGAGACGATCAC  811

Qy 783   GATCATTTACTTACAGATGTCCAAGATGTAAAGTCAGATGTTGCCATTTGGTTATGAG  842
         | ||||||||| || | ||| ||||||||||  |||||||||||||||||||||| |
Db 812   GCTCATTTACCTGCAATGAATTCCAAGATGTACAATCAGATGTTGCCATTTGGTTATAAA  871

Qy 843   TAAAAGTCATTTGTCTTCTTTGATAGTGTATGTGTATCTGAAAATATATACTAATACT  902
         ||||| |||||| | |||| ||||| |  || | || |||||||| | || | | | |
Db 872   GAAAATTGATTTTGCTCATTTTATGACAAACTTATACATC-------------------  912

Qy 903   AATGTGTACTTCTAAAATATAGCTTGTGTATAAGAAGAGATGATTTCATTTT-----TA  956
          |||| |||||||||||| || | ||| ||| |||| ||| | | ||   ||     ||
Db 913   ------TGCTTCTAACATATCGCATGTTTATGTTAAGATTTGGTCCCATCCTTTAAACTG  966

Qy 957   AAATACACCATGCAAATACATATTAAAATGTAAGAACTTTTTATTATTATACTAAAATAATT 1016
         ||||| | ||||| |||| ||||| ||||||||||| |||||||| ||||| |||||||||
Db 967   AAATATGTCATGTGAAATTATTTTAAAAATGTAAAAAACAAAACTTTCTGCTAACAAAATA 1026

Qy 1017  CATCATCTATCTTCCGAAATATTTTATGAAAATCTATCTGATATTCTATTCTAATAAAAT 1076
         |||  || |||   |||| ||||| | ||||| ||||| |||||||| ||  ||||| ||
Db 1027  CATACAGTATCTGCCAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTATAA 1086

Qy 1077  TC 1078
         ||
Db 1087  TC 1088
```

RESULT 5
US-09-894-749-1
; Sequence 1, Application US/09894749
; Patent No. US2020081683A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; APPLICANT: Yowe, David
; TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
; FILE REFERENCE: 5800-19, 035800/174680
; CURRENT APPLICATION NUMBER: US/09/894,749
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/244,314
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (160)..(867)
US-09-894-749-1

Query Match   51.2%;   Score 596;   DB 9;   Length 2217;
Best Local Similarity   75.8%;   Pred. No. 1.1e-118;
Matches 791;   Conservative   0;   Mismatches 220;   Indels   33;   Gaps   3;

```
Qy  41   TTTATTCAGGATGTTTCCTATGAAATAGCATTCATCTGTGGGAGAGAGGAAGACTAAGG  100
         |||||||  ||| || |||||| |||||| ||| || ||||| ||  ||||||| ||||
Db  69   TTATTCTACTACATGTGATATGTGGATTAATAATGAACTAGGGGAGGATGTAAT  128

Qy 101   AAATCTGACATCTGTGGTCACTGGGACAGAATATGGATTATGTCACTGGTTTCTTCTCT  160
         ||||  ||||| || | ||   ||| |||||| ||  | ||||||||||||||| ||||
Db 129   AAATTAGACATTCTTCAT--TTTAGAGAAGATGGAAACAACATTGCTTTCTTTTTCT  186

Qy 161   CAATTAAATATGTGAATCAAAAGAAGAAAACTTTTTCAAACTAATGCATGGGTCAGGG  220
         ||| |||||||||||||||||||| | ||||||||||||  |||| |||||| ||||
Db 187   CAAATAAATATGTGAATCAAAAGAAAAAAAACTTTTTCAAGTTAATACATGGTTCAGGA  246
```

; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 847
; LENGTH: 848
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: wherein any n is one of a or t or g or c
US-09-867-550-847

Query Match   48.2%;   Score 560.6; DB 9;   Length 848;
Best Local Similarity 82.0%;   Pred. 3.1e-111;
Matches 670; Conservative   0; Mismatches 144; Indels   3; Gaps   2;

```
Qy      1 TTTTTGTAAGAAAAATGTGAGGAAAGATTCGGATAGCGCTTTATTC-AGGATGTTTTCC 59
            ||||| ||||||||| |   ||  |      || |||||||||||| ||||||||
Db     34 TTCTTTTTGTAAACATTACTGTAAGAGTTGTGATAACTTTTTATTCTACTATGTATATG 93

Qy     60 TATGAAATAGACATTCATCTGTGGGAGAGAAGAACTGACATCTGTTGGT 119
           |||||||||| |||  |||| || |||| |  |  ||||| ||| |||||
Db     94 TATGGAATAGTATTAATAAACATAGGGAGGAGGATGTAATAAATTAGACATCTCTTCAT 153

Qy    120 CACTGGGACAGAATATGGATAGTCACTGGTTTTCTTCTCTCAATTAAATATGTGAAT 179
           ||||| ||| ||  ||||| |||||||  ||| |  ||||   ||| || ||||||||
Db    154 T--TTAGGAGAAGATGGAAACAACATTGCTTTTCTTCTTTCTCAAATAAATATGTGAAT 211

Qy    180 CAAAAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGGAAGAAGAAACAAGCATCG 239
           |||||||||||||||||  ||||||||| |  |||  | |||||| ||||| |||  ||
Db    212 CAAAAGAAAAAACTTTTTCAAGTTAATACATGGTTCAGGAAATGGTTCAGAGAACAAGCAAAG 271

Qy    240 AGGCCAAAATCAGAGCGAAAGAAAAAAGGAATAGACTAAGTCTCTTCCTACAGAGGCCTG 299
           |||||||| |||||||| || |  || |||| ||||  |||||| ||||| | |||||| |
Db    272 AAGCCAAAATCAGAGCTAAGAGAAAAAGAAAATAGACTTAAGTCTTCTGTGCGAAACCTG 331

Qy    300 ACTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCT 359
           ||| |||  ||| | || |   ||  |||||| ||   ||||||||||||| |||||| |
Db    332 AGTTTCATGAAGACACACCCGCTCAGTAGATCGGGGACTTGGCCAAAGAAACAAGAGTCT 391

Qy    360 CTCCTGAAGAAGCAGCAGTGAAATGGGCTGAATCATTTTGACAAATGCTCTCCATAGAGATG 419
           |  |||||| |||| |  || ||| | ||  |||| ||||||| || |   || |||||||
Db    392 CCCCTGAAGAGAGCCAGTGAAATGGGTGAATCATTTGACAAACTGCTTTCCCATAGAGATG 451

Qy    420 GAGTGGATGCTTTCACCAGATTCTTAAACTGAAATGAGGAGAACATTGAATTTT 479
           |||| || || |||| |||| | ||||||| || ||  |  ||| |  ||||||||
Db    452 GACTAGAGGGCTTTCACCAGATTCTTAAAACTGAATTCAGTGAAGAAATATTGAATTTT 511

Qy    480 GGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAA 539
           ||| ||||  || |||||| |||||| |||| |  || ||  ||||||||   |||||||
Db    512 GGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGAGACCTCAACAAATTCACCTTAAAGCAA 571

Qy    540 AGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTTC 599
           || | ||| ||||||||| || ||||| || ||| |   |||||||| |||||||| |
Db    572 AAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAAGAGGTAACCTTGATTTTC 631

Qy    600 ATACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCACTCCTCCACAGTTTTGATACGG 659
           || || |||| | |||| || ||||| |||| |    |  |  || ||||| |||| | |
Db    632 ACACAAAGAAGAGTCATTAACAACAGCATCCCACTCCAGCTTAAACGCTTTCCACAGTTTTGATGCTG 691

Qy    660 CACAAAGCAGAGTGTACCAGCTCATGGAAATCTG 719
           |||||||||||| ||| |||| | ||||||| |||
Db    692 CACAAAGCAGAGTGTATCAGCTCATGGAACAAGACAGAAGAGTTATACACGTTTCTGAAATCTG 751

Qy    720 AGACCTACTTACATTTGATAGAAGGAAGACCTCAGAGACCAACAACAAACCTTAGGAGACGAT 779
           ||  | || |  || |||  | |||||||||| |||||||||||| ||| ||||| ||||||
Db    752 ACATCTATTAGACTTGATGGAGGAAGGAAGACCTCAGAGACCAACAACAAATCTTAGGAGACGAT 811
```

```
Qy    221 AAAGAAGAAACAAGCATCGAGCGCAAAATCAGAGCGCGAAAGAAAAAGGAATGACTAAGT 280
            |||||||||||||||  ||| |||| ||||  |||||   || | |||| ||||||||||
Db    247 AAAGAAGAAACAAGCAAAGAGAAGCCAAAATCAGAGCTAAGAGAAAAATAGGACTAAGT 306

Qy    281 CTTCCTCCACAGAGGCCTGACTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCTCTTG 340
           ||| | |||  ||  | |||||  | |||||| |||   || ||| |||||| || ||||
Db    307 CTTCTTGTGCAGAAACCTGAGTTTCATGGAGAGACTCAAGCCAGTAGATCTGGGCACTTG 366

Qy    341 GCCAAAGAAACAAGAGTCTCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTGACAAA 400
           || |||||||| ||||||| || || |||| |||||| ||| | || || ||||||||||
Db    367 GCCAAAGAAACAAGAGTCTCCCCTGAAGAGAGCCAGTGAAATGGGTGAATCATTGACAAA 426

Qy    401 TTGCTCTCTCATAGAGATGGAGTGGATGCTTTTCACCAGATTCTTAAACTGAATTCAGT 460
           |  || |||  ||  | ||| || |  ||||||||||||||||||||||| |||||||||
Db    427 CTGCTTTCCCATAGAGATGGACTAGAGGGCTTTCACCAGATTCTTAAAACTGAATTCAGT 486

Qy    461 GAGGAGAACATTGAATTTGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAA 520
           || | |  ||||||||| |||| ||||  || |||||||||||| |||| |  || ||||
Db    487 GAAGAAAATATTGAATTTGGATAGCCTGTGAAGATTTCAAGAAAAGCAAGGAGACCTCAA 546

Qy    521 CAAATCATCCTAAAAGCAATCAATCAAAGCAGAATCATTCAGAATGAATGACCCCAAA 580
           |||| |||   ||||||||| | |||||||||||| || ||||| || ||| ||||||||
Db    547 CAAATTCACCTTAAAGCAAAGCAATATATGAGAAATTTATACAGACTGATGCCCCAAAA 606

Qy    581 GAGGTTAACATTGATTTCATACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACT 640
           |||| |||| |||||||||||| | || ||||  | || |||| |||| |  | || |  |
Db    607 GAGGTTAACCTTGATTTTCATACAAAGAAGAGTCATTAACAACAGCATCCCACTCCAACCTACC 666

Qy    641 CTCCACAGTTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTAT 700
           ||||||||||||||| | |||||||||||||| ||| |||| | |||||| || |||||||
Db    667 CTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGAATGAATCAGTTATA 726

Qy    701 AAACGCTTTTGAAATCTGAGAACCTACTACTTGATAGAAGGAAGACCTCAGAGACCA 760
           | ||| || |  || |||||  || || |  |||||  | |||||||||||||||||||
Db    727 ACACGTTTCTGAAATCTGACATCTATTAGACTTGATGGAGGAAGGAAGACCTCAGAGACCA 786

Qy    761 ACAAACCTTAGGAGACGACGATCACGATCATTACTTACAATGATTTCCAAGATGTAAAGTCA 820
           ||||| || |||||| ||| |||| | |||   || ||||||  ||| ||||| |||| || |
Db    787 ACAAATCTTAGGAGACGACGATCACGCGTTCATTACCTGCAATCATGAATTCCAAGATGTACAATCA 846

Qy    821 GATGTTGCCATTTGGTTATGAGTAAAGTCATTTGTCTTCTTTGATAGTGTATGTGTAT 880
           |||||||||||| |  ||| |||||||| |||| || |  || |  ||||  |||| |
Db    847 GATGTTGCCATTTGGTTATAAGAAAAAATTGATTTTGCTCATTTTTATGCACAAACTTATAC 906

Qy    881 ATCTAAAATATATATACTAATGTGTACTTCTAAAATAATACATATTGTGTATAAGAAGA 940
              ||||||| ||||| |||||| |||| | |  |||||| ||||| || ||||||||||
Db    907 ATC------------------TGCTTCAACATATCGCATGTTTATGTTAAGA 941

Qy    941 GATGATTTCATTTT------TAAAATACACCATGCAAATACATATTAAATGTAAGAACTT 994
            ||| || ||||||      | |||| |   || ||  |||  ||||||  ||| ||||
Db    942 TTTGGTCCCATCCTTAAACTGAAATATGTCATGTGAAATTATTTAAAAATGTAAAAAC 1001

Qy    995 TTTATATTATACTAAATAATCATCATCATCTCCGAAAATATTTATGAAAATCTATC 1054
           |||| | ||| || ||   | |  | | |||| | |||| | ||| | ||||||||
Db   1002 AAAACTTTCTGCTAACAAATACATACAGTAGATCTCGCCAGTATATTCTGTAAAACCTTCTA 1061

Qy   1055 TGATATTCTATTCTAATAAAATTC 1078
           || |  |||| || |||  |||||
Db   1062 TTTGATGTCATTCCATTTATAATC 1085
```

RESULT 6
US-09-867-550-847
; Sequence 847, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US2002008220641e1 Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby

```
Qy   780  CACGATCATTACTTACCAAGATGATTTCCAAGATGTAAA  816
               ||||||||||||| ||||||  |||||||||||||
Db   812  CACGCTCATTACCTGCAATGAATTCCAAGATGTACA     848
```

RESULT 7
US-09-833-381-1010
; Sequence 1010, Application US/09833381
; Patent No. US2002013209OA1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090Alel Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1010
; LENGTH: 736
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(736)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1010

Query Match   26.2%;   Score 305.4; DB 9; Length 736;
Best Local Similarity 86.8%; Pred. No. 5e-56;
Matches 336; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

```
Qy   455  TTCAGTGAGGAGAACATTGAATTTGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAA  514
               |||||||| |||||||||||||| || ||  |||||||||| ||||||||| |||||||
Db     4  TTCAGTGAAGAAATATTGAATTTTGATAGCGTGTTGAAGATTTCAAGAAAAGCAAGGGA   63

Qy   515  CCTCAACAAATCATCCTAAAAGCAATCTATGAGAAATCATTCAGAATGATGCC  574
               ||||||||||| |||| ||||||||   |||||||| ||||   ||||||||||
Db    64  CCTCAACAAATTCACCTTAAAGCAAAGCAATATTGAGAAATTATACGACTGATGCC  123

Qy   575  CCCAAAGAGGTTAACATTGATTTTCATACTAAAGAAGTAATTGCTAAGGACATCGCCCAG  634
               || |||||||||||| ||||||||| || ||||||||| ||| ||  |||||||||||||
Db   124  CCAAAAGAGGTTAACCTTGATTTTCACACAAAAGAAGTCATTACAACAGCAGATCACTCAA  183

Qy   635  CCCACTCTCCACAGTTTTGATACGGACACAAAGCAGACGAGTGTACCCAGCTCATGGAACATGAC  694
               || || |||||||||||||| ||| || ||||||||| ||||||| | ||||||||| ||||| |||||
Db   184  CCTACCCCTCCACAGTTTTGATGCTGCACAAAGCAGAGAGTGTATCAGCTCATGCAATGAACAGAC  243

Qy   695  AGTTATAAACGCTTTTGAAATCTGAGACCTACCATTGATGAGAAGGAAGACCTCAG  754
               ||||||| |||||||||||||||| |||| ||||| |||||  |||||||||||||
Db   244  AGTTATACACGTTTTGAAATCTGACATCATTTAGACTTGATGGAAGGAAGACCTCAG  303

Qy   755  AGACCAACAAACCTTAGGAGACGATCACGATCATCACGCGCTCATTCCAAGATGTA  814
               ||||||||||  |||||||||||||||||||  || || |||||||||||||||||
Db   304  AGACCAACAACAAATCTTAGGAGACGATCACGCGCTCACGCGCTCATTCCAAGATGTA  363

Qy   815  AAGTCAGATGTTGCCATTTGGTTATGA  841
               || |||||||||||||||||| ||||||
Db   364  CAATCAGATGTTGCCATTTGGTTATAA  390
```

RESULT 8
US-10-258-371B-6
; Sequence 6, Application US/10258371B
; Publication No. US20040067903A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS-GAGNON, Alison
; APPLICANT: MURRAY, David L
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING,
; TITLE OF INVENTION: RGS18, AND USES THEREOF
; FILE REFERENCE: A3656 US PCT
; CURRENT APPLICATION NUMBER: US/10/258,371B
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: GB001883.334
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US60/200,786
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-258-371B-6

Query Match   16.5%;   Score 192.2; DB 13; Length 1486;
Best Local Similarity 68.4%; Pred. No. 2.2e-31;
Matches 311; Conservative 0; Mismatches 113; Indels 31; Gaps   2;

```
Qy   630  CCCAGCCCACTCTCCACAGTTTTGATACGGCACAAGCAGAGAGTGTACCAGCTCATGGAAC  689
               |||| | ||| |||||||||||||| |||| || |||| || ||||||| |||||||||||
Db     1  CTCAACCTACCCCTCCACAGTTTTGATGCTGCACAAAGCAGAGAGTGTATCAGCTCATGGAAC   60

Qy   690  ATGACAGTTATAAAACGCTTTTGAAATCTGAGACCTACCATTTGATAGAGGAGAC  749
               | || |||||||| | ||| |||||||||||||||||   |||| || |||||||||
Db    61  AAGACAGTTATACACGTTTTTCTGAAATCTGAGACCTATTTGAGTCGTGATGGGAAGGAAGAC  120

Qy   750  CTCGAGAGACCAACAAACCTTAGGAGACGATCACGATCATTACTTACAGATTCCAAG  809
               ||||||||||||||||  |||||||||||||||||| |  |||||  ||||||||||
Db   121  CTCGAGAGACCAACAACAAATCTTAGGAGACGATCACGCGCTCATTCCAAG  180

Qy   810  ATGTAAAGTCAGATGTTGCCATTTGGTTATGATTTATGAGTAAAAGTCATTTGTCTTCTTTTGATAG  869
               |||| |||| ||||||||| ||||||||||  |||| |||| |||| ||| |  | |  | |||||
Db   181  ATGTACAATCAGATGTTGCCATTTGGTTATAAAGAAAATTGATTTTGCTCATTTTTATGA  240

Qy   870  TGTATGTGTATATCTAAAATATATACTAATATCAATGTGTACTTCTAAAATATAGCTTGT  929
               ||| || | | | | || |  ||  || |
Db   241  CAAACTTATCATC------------------TGCTTCTAACATATCGCATGT  275

Qy   930  GTATAAGAAGAGATGAATTCATTT-----TAAAATACACCATGCAAATACATATTAAA  983
               |||| | |||| |||||| |||||||     ||| | ||| | ||||||| | ||||||
Db   276  TTATGTTAAGATTTGGTCCCATCCTTAAACTGAAATATGTCATGTGAAATTATTTTAAA  335

Qy   984  TGTAAGAACTTTTATATTCATACTAAAGAATAATTCATCATCTACTCTTCCGAAAATAATTTAT  1043
               |||||||| |||| | ||| | ||| ||||| |||  | || |||||| |||||||||
Db   336  AATGTAAAAACAAAACTTTCTGCTAACAAAATACACTACAGTATCCTGCCAGTATATTCTGT  395

Qy   1044  GAAAATCATCTGATATTCTATTCTAATAAAATTC  1078
                | ||||||||| |||| ||| || ||| ||||||
Db    396  AAAACCTTCTCATTTGATGTCATTCCATTTATATATAATC  430
```

RESULT 9
US-10-258-371B-11
; Sequence 11, Application US/10258371B
; Publication No. US20040067903A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS-GAGNON, Alison
; APPLICANT: MURRAY, David L
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING, AND USES THEREOF
; TITLE OF INVENTION: RGS18, AND USES THEREOF
; FILE REFERENCE: A3656 US PCT
; CURRENT APPLICATION NUMBER: US/10/258,371B
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: GB001883.334
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US60/200,786
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-258-371B-11

Query Match   15.3%;   Score 177.6; DB 13; Length 241;
Best Local Similarity 83.8%; Pred. No. 1.3e-28;
Matches   201; Conservative   0; Mismatches   39; Indels   0; Gaps   0;

Qy 461 GAGGAGAACATTGAATTTGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGAACCTCAA 520
            |||||| || ||||| |||||||||||||| ||||||||||| ||| ||| ||||||
Db   1 GAGGAAAATCTGGAGTTCTGGATAGCCTGTGAAGATTTCAAGAAAGCAAGGGACTTCAA 60

Qy 521 CAAATCATCCTAAAGCAAGGCAATCATCTATGAGAAATTCATTCAGAATGATGCCCCAAA 580
            ||||| | ||| ||||||| |||| |||||||||||| | | |||| |||||||||||
Db   61 CAAATTCACCTTAAAGCAAAGCAATATATGAGAAAT TATACAGACTGATGCCCCAAA 120

Qy 581 GAGGTTAACATTGATTTCATACTAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACT 640
            |||||||| |||||||||| | ||||||| || || |||||| | ||| || ||
Db 121 GAGGTTAACCTTGATTTCACACAAAGAAGTCATTACAAACAGCATCACTCAACTACC 180

Qy 641 CTCCACAGTTTGATACGGCACAAAGCAGAGTGTACCAGTCATGGAACATGACAGTTAT 700
            |||||||||||| | | ||||||||||||||| || |||||||||| | ||||||| ||
Db 181 CTCCACAGTTTGTGCACAAAGCAGAGTGTATCAGTCATGGAAAAACGACACAGCTAT 240

RESULT 10
US-10-342-887-839
; Sequence 839, Application US/10342887
; Publication No.US20040058340A1
; GENERAL INFORMATION:
;   APPLICANT: Dai, Hongyue
;   APPLICANT: He, Yudong
;   APPLICANT: Linsley, Peter S.
;   APPLICANT: Mao, Mao
;   APPLICANT: Roberts, Christopher J.
;   APPLICANT: Van 't Veer, Laura Johanna
;   APPLICANT: Van de Vijver, Marc J.
;   APPLICANT: Bernards, Rene
;   TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;   FILE REFERENCE: 9301-188-999
;   CURRENT APPLICATION NUMBER: US/10/342,887
;   CURRENT FILING DATE: 2003-01-15
;   PRIOR APPLICATION NUMBER: 60/298,918
;   PRIOR FILING DATE: 2001-06-18
;   PRIOR APPLICATION NUMBER: 60/380,710
;   PRIOR FILING DATE: 2002-05-14
;   PRIOR APPLICATION NUMBER: 10/172,118
;   PRIOR FILING DATE: 2002-06-14
;   NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 839
;   LENGTH: 1345
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-342-887-839

Query Match   13.1%;   Score 152.4; DB 13; Length 1345;
Best Local Similarity 62.7%; Pred. No. 8.7e-23;
Matches   237; Conservative   0; Mismatches 141; Indels   0; Gaps   0;

Qy 359 TCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGAT 418
            |||||||| |||||||| | ||| ||||||||||||||||| ||||| | ||| |||
Db 249 TCTCCTGAGGAAGCACAGCGCTGTGGTCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy 419 GGAGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 478
            || | | |||| || | | | ||||| ||| | ||||| || || |||| ||||||||
Db 309 GGTCTTCTGCTCATTCAGGGCTTTTTTAAAGTGCGGATTCTGTGAAGAAAATATTGAATTC 368

Qy 479 TGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGAACCTCAACAAATCATCCTAAAAGCA 538
            ||| | ||||||||||||||||| || | ||| ||||| ||||| || || ||||||
Db 369 TGGCTGGCCTGTGAAGACTTCAAAAAAACCAAATCACCCCAAAAGCTGTCTCAAAAGCA 428

Qy 539 AAGGCAATCTATGAGAAATTCATTCAGAAATGCTGAAATCAGTGAGGAGGTTAACATTGATTTT 598
            ||| ||| | |||| |||||| |||| | | |||| || || |||| ||||||||
Db 429 AGGAAAATATATACTGACTTCATAGAAAAAGAAGCTCCAAAAGAGATAAACATAGATTTT 488

Qy 599 CATACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCGCCACTCTCCACAGTTTGATACG 658

Db 489 CAAACCAAAACTCTGATTGCCCAGAATATACAAGAAGCTACAAGTGGCTGCTTTACAACT 548
            |||| | |||||||||||||||||| |||||||||||||||||||||||||| |||| ||
Qy 659 GCACAAAGCAGAGTGTACCAGTCATGGAACATGCAGTTATAAACGCTTATAAACCTTATCCTCGTTTTGGAAATCT 718
            || |||| ||||||| |||||| |||||| | | |||||||| ||| || |||||||||||||| |||| | |
Db 549 GCCCAGAAAAAGGGTATACAGCTTGATGGAGAACAACTTCTTATCCTCGTTTTCTGGAGTCA 608

Qy 719 GAGACCTACTTACATTG 736
            |||||| |||| ||||
Db 609 GAATTCTTACCAGGACTTG 626

RESULT 11
US-10-172-118-839
; Sequence 839, Application US/10172118
; Publication No.US20030224374A1
; GENERAL INFORMATION:
;   APPLICANT: Dai, Hongyue
;   APPLICANT: He, Yudong
;   APPLICANT: Linsley, Peter
;   APPLICANT: Mao, Mao
;   APPLICANT: Roberts, Chris
;   APPLICANT: Van 't Veer, Laura
;   APPLICANT: Van de Vijver, Marc
;   APPLICANT: Bernards, Rene
;   TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;   FILE REFERENCE: 9301-175-999
;   CURRENT APPLICATION NUMBER: US/10/172,118
;   CURRENT FILING DATE: 2002-06-14
;   PRIOR APPLICATION NUMBER: 60/380,770
;   PRIOR FILING DATE: 2002-05-14
;   NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 839
;   LENGTH: 1345
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: NM_002923
;   DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-839

Query Match   13.1%;   Score 152.4; DB 13; Length 1345;
Best Local Similarity 62.7%; Pred. No. 8.7e-23;
Matches   237; Conservative   0; Mismatches 141; Indels   0; Gaps   0;

Qy 359 TCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGAT 418
            |||||||| |||||||| | ||| ||||||||||||||||| ||||| | ||| |||
Db 249 TCTCCTGAGGAAGCACAGCGCTGTGGTCGAGAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy 419 GGAGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 478
            || | | |||| || | | | ||||| ||| | ||||| || || |||| ||||||||
Db 309 GGTCTTCTGCTCATTCAGGGCTTTTTTAAAGTGCGGATTCTGTGAAGAAAATATTGAATTC 368

Qy 479 TGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGAACCTCAACAAATCATCCTAAAAGCA 538
            ||| | ||||||||||||||||| || | ||| ||||| ||||| || || ||||||
Db 369 TGGCTGGCCTGTGAAGACTTCAAAAAAACCAAATCACCCCAAAAGCTGTCTCAAAAGCA 428

Qy 539 AAGGCAATCATGAGAAATTCATTCAGAATGATGCTGAAATCAGTGAGGAGGTTAACATTGATTTT 598
            ||| ||| | |||| |||||| |||| | | |||| || || |||| ||||||||
Db 429 AGGAAAATATATACTGACTTCATAGAAAAAGAAGCTCCAAAAGAGATAAACATAGATTTT 488

Qy 599 CATACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCGCCACTCTCCACAGTTTGATACG 658
            || ||||| |||||| || |||| | || ||| ||||| ||||||||||||||||| ||||
Db 549 GCCCAGAAAAAGGGTATACAGCTTGATGGAGAACAACTTCTTATCCTCGTTTTCTGGAGTCA 608

Qy 719 GAGACCTACTTACATTG 736
            |||||| |||| ||||
Db 609 GAATTCTTACCAGGACTTG 626

RESULT 12
US-10-305-720-1278
; Sequence 1278, Application US/10305720
; Publication No. US2004010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expressio
; FILE REFERENCE: PA-0002:1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1278
;   LENGTH: 1345
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: GenBank ID No. US2004010136A1 g292054
US-10-305-720-1278

Query Match   13.1%;   DB 16;   Score 152.4;   Length 1345;
Best Local Similarity 62.7%;   Pred. No. 8.7e-23;
Matches 237;   Conservative   0;   Mismatches 141;   Indels   0;   Gaps   0;

Qy   359 TCTCCTGAAGAAGCAGTGAAATGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGAT 418
          |||||| |||| |||||| ||||||| ||
Db   249 TCTCCTGAGGAAGCACAGCTGTCGTCGTCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy   419 GGAGTGGATGCTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 478
          |||| ||| |||| |||||| ||| |||||||| ||||
Db   309 GGTCTTGCTGCATTCAGGGCGTTTTTAAAGTCGGAATTCTGTGAAGAAAATATTGAATTC 368

Qy   479 TGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAAACAAATCATCCTAAAAGCA 538
          ||||| ||||||||||||||||| ||||||||||| |||||| |||||
Db   369 TGGCTCGGCCTGTGAAGACTTCAAGAAGACTTCAAGAAATCACCCCAAAAAGCGTGCTCCTCAAAAGCA 428

Qy   539 AAGGCAATCTATGAGAAAATCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTT 598
          ||| ||||| ||||||||||||| ||||||| |||||
Db   429 AGGAAAATATATACTGACTTCAGTGAACATGACAGTTATAAACATAGATTTT 488

Qy   599 CATACTAAAGAAGTAAATGCTAGAGACATCGCCCAGCCACTCTCCACAGTTTGATACG 658
          ||||| |||||| || ||||||||||||| |||||| ||||||||
Db   489 CAAACCAAAACTCTGATTGCCCAGACAAGAAGCTACAAGTGGCTGCTTTACAACT 548

Qy   659 GCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTGAAATCT 718
          |||||| |||||||||| |||||||||||||||||| |||||||
Db   549 GCCCAGAAAAGGGTATACAGCTTGATGGAGAACAACTCTTATCCTGTTTCTTGGAGTCA 608

Qy   719 GAGACCTACTTACATTTG 736
          ||||||||||||||||
Db   609 GAATTCTACCAGGACTTG 626

RESULT 13
US-10-641-643-1243
; Sequence 1243, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
;   Susan G. Stuart
;   Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
;   GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/641,643
;   FILING DATE: 14-Aug-2003
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: <Unknown>
;   FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1243:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1345 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   LIBRARY: GENBANK
;   CLONE: g292054
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1243 :
US-10-641-643-1243

Query Match   13.1%;   DB 17;   Score 152.4;   Length 1345;
Best Local Similarity 62.7%;   Pred. No. 8.7e-23;
Matches 237;   Conservative   0;   Mismatches 141;   Indels   0;   Gaps   0;

Qy   359 TCTCCTGAAGAAGCAGTGAAATGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGAT 418
          |||||| |||| |||||| ||||||| ||
Db   249 TCTCCTGAGGAAGCACAGCTGTCGTCGTCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy   419 GGAGTGGATGCTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 478
          |||| ||| |||| |||||| ||| |||||||| ||||
Db   309 GGTCTTGCTGCATTCAGGGCGTTTTTAAAGTCGGAATTCTGTGAAGAAAATATTGAATTC 368

Qy   479 TGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAAACAAATCATCCTAAAAGCA 538
          ||||| ||||||||||||||||| ||||||||||| |||||| |||||
Db   369 TGGCTCGGCCTGTGAAGACTTCAAGAAGACTTCAAGAAATCACCCCAAAAAGCGTGCTCCTCAAAAGCA 428

Qy   539 AAGGCAATCATGAGAAAATCATTCAGAATGATGCCCCAAGAGCATGCGCCCAGCCACTCTCCACAGTTTTGTATACG 598
          || ||||| ||||| |||||||||| ||||||| |||||
Db   429 AGGAAAATATATACTGACTTCAGTGAACATGACAGTTATAAACATAGATTTT 488

Qy   599 CATACTAAAGAAGTAATTGCTAAGAGCATGGAACATGACAGTTTTGTGATACG 658
          ||| ||||||||| || |||||||||||||| |||||| ||||||
Db   489 CAAACCAAAACTCTGATTGCCCAGACAAGAAGCTACAAGTGGCTGCTTTACAACT 548

Qy   599 GCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAAATCT 718
          |||||| |||||||||| |||||||||||||||||| |||||||
Db   549 GCCCAGAAAAGGGTATACAGCTTGATGGAGAACAACTCTTATCCTGTTTCTTGGAGTCA 608

Qy   719 GAGACCTACTTACATTTG 736
          ||||||||||||||||
Db   609 GAATTCTACCAGGACTTG 626

RESULT 14
US-09-925-300-567
; Sequence 567, Application US/09925300
; Patent No. US2002015181A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben

```
;    FEATURE:
;      NAME/KEY: misc_feature
;      OTHER INFORMATION: Incyte ID No. US20030175704A1 989992.12
US-09-971-429B-32

     Query Match         13.1%;    Score 152.4;  DB 10;   Length 1381;
     Best Local Similarity  62.7%;  Pred. No. 8.9e-23;
     Matches    237;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy   359 TCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGAT 418
         |||||  ||||||  |||||||  ||||||| |  | |||  |||||  |||||  ||||
Db   262 TCTCCTGAGGAAGCACCAGCTGTGGTCGGTCAGCAGCATTTGACGAGCTGCTAGCCAGCAAATAT 321

Qy   419 GGAGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 478
         |||| ||  ||| ||  ||  ||| |||  |||||||  |  ||||||||| |||||||
Db   322 GGTCTTGCTGCATTCAGGGCTTTTTAAAGTCGGAATTCTGTGAAGAAAATATTGAATTC 381

Qy   479 TGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAGAACCTCAACAAATCATCCTAAAAGCA 538
         |||||  ||||||||||||||||||  ||| |||||||| |||| |||||||||||||||||
Db   382 TGGCTGGCCTGTGAAGACTTCAAAAAACCAAATCACCCCAAAAAAGCTGTCCTCAAAAGCA 441

Qy   539 AAGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTT 598
         ||  |||| | ||||||| ||  ||||| |  | ||||  |||||||  ||||||||||
Db   442 AGGAAAATATATACTGACTTCATAGAAAAGGAAGCTCCAAAAGAGATAAACATAGATTTT 501

Qy   599 CATACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGATACG 658
         || |||||||||||  |||  | |||||  | || ||||  | | | || |||||| ||
Db   502 CAAACCAAAACTCTGATTGCCCAGAATATACAAGAGCTACAAGTGGCTGCTTTACAACT 561

Qy   659 GCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAAATCT 718
         ||| ||||||||  |||  |  | |||  || |  |  ||||  ||||| |  |||||||
Db   562 GCCCAGAAAAGGGTATACAGCTTGATGGAGAACAACTCTTATCCTCGTTTCTCGGAGTCA 621

Qy   719 GAGACCTACTTACATTTG 736
         ||||  | |   |||||
Db   622 GAATTCTACCAGGACTTG 639

Search completed: August 20, 2004, 16:40:48
Job time : 546.613 secs
```

```
;    TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;    FILE REFERENCE: PA101
;    CURRENT APPLICATION NUMBER: US/09/925,300
;    CURRENT FILING DATE: 2001-08-10
;    PRIOR APPLICATION NUMBER: PCT/US00/05988
;    PRIOR FILING DATE: 2000-03-08
;    PRIOR APPLICATION NUMBER: 60/124,270
;    PRIOR FILING DATE: 1999-03-12
;    NUMBER OF SEQ ID NOS: 1890
;    SOFTWARE: PatentIn Ver. 2.0
;    SEQ ID NO 567
;    LENGTH: 1364
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;      NAME/KEY: misc_feature
;      LOCATION: (1362)
;      OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-567

     Query Match         13.1%;    Score 152.4;  DB 9;   Length 1364;
     Best Local Similarity  62.7%;  Pred. No. 8.8e-23;
     Matches    237;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy   359 TCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGAT 418
         |||||  ||||||  |||||||  ||||||| |  | |||  |||||  |||||  ||||
Db   246 TCTCCTGAGGAAGCACCAGCTGTGGTCGGTCAGCAGCATTTGACGAGCTGCTAGCCAGCAAATAT 305

Qy   419 GGAGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 478
         |||| ||  ||| ||  ||  ||| |||  |||||||  |  ||||||||| |||||||
Db   306 GGTCTTGCTGCATTCAGGGCTTTTTAAAGTCGGAATTCTGTGAAGAAAATATTGAATTC 365

Qy   479 TGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAGAACCTCAACAAATCATCCTAAAAGCA 538
         |||||  ||||||||||||||||||  ||| |||||||| |||| |||||||||||||||||
Db   366 TGGCTGGCCTGTGAAGACTTCAAAAAACCAAATCACCCCAAAAAAGCTGTCCTCAAAAGCA 425

Qy   539 AAGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTT 598
         ||  |||| | ||||||| ||  ||||| |  | ||||  |||||||  ||||||||||
Db   426 AGGAAAATATATACTGACTTCATAGAAAAGGAAGCTCCAAAAGAGATAAACATAGATTTT 485

Qy   599 CATACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGATACG 658
         || |||||||||||  |||  | |||||  | || ||||  | | | || |||||| ||
Db   486 CAAACCAAAACTCTGATTGCCCAGAATATACAAGAGCTACAAGTGGCTGCTTTACAACT 545

Qy   659 GCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAAATCT 718
         ||| ||||||||  |||  |  | |||  || |  |  ||||  ||||| |  |||||||
Db   546 GCCCAGAAAAGGGTATACAGCTTGATGGAGAACAACTCTTATCCTCGTTTCTCGGAGTCA 605

Qy   719 GAGACCTACTTACATTTG 736
         ||||  | |   |||||
Db   606 GAATTCTACCAGGACTTG 623

RESULT 15
US-09-971-429B-32
;    Sequence 32, Application US/09971429B
;    Publication No. US2003017570T4A1
;    GENERAL INFORMATION:
;    APPLICANT: Lasek, Amy K. W.
;    APPLICANT: Shyjan, Andrew W.
;    APPLICANT: Turner, Christopher M.
;    TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER
;    FILE REFERENCE: PA-0040 US
;    CURRENT APPLICATION NUMBER: US/09/971,429B
;    CURRENT FILING DATE: 2001-10-04
;    PRIOR APPLICATION NUMBER: 60/239,024
;    PRIOR FILING DATE: 2000-04-10
;    NUMBER OF SEQ ID NOS: 56
;    SOFTWARE: PERL Program
;    SEQ ID NO 32
;    LENGTH: 1381
;    TYPE: DNA
;    ORGANISM: Homo sapiens
```

```
                    GenCore version 5.1.6
         Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      August 20, 2004, 06:15:45 ; Search time 462.364 Seconds
                                         (without alignments)
                                         7750.315 Million cell updates/sec

Title:       US-09-894-749-3
Perfect score: 1164
Sequence:    1 tttttgtaagaaaatctga.........gtactagtggatccgagctc 1164

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched:    5021620 seqs, 1539289099 residues

Total number of hits satisfying chosen parameters:    10043240

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Pending_Patents_NA_New:*
  1: /cgn2_6/ptcdata/2/pna/PCT_NEW_COMB.seq:*
  2: /cgn2_6/pcodata/2/pna/US06_NEW_COMB.seq:*
  3: /cgn2_6/pcodata/2/pna/US07_NEW_COMB.seq:*
  4: /cgn2_6/pcodata/2/pna/US08_NEW_COMB.seq:*
  5: /cgn2_6/pcodata/2/pna/US09_NEW_COMB.seq:*
  6: /cgn2_6/pcodata/2/pna/US10_NEW_COMB.seq:*
  7: /cgn2_6/pcodata/2/pna/US10_NEW_COMB.seq2:*
  8: /cgn2_6/pcodata/2/pna/US10_NEW_COMB.seq3:*
  9: /cgn2_6/pcodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	602.6	51.8	1884	8	US-10-115-635-133	Sequence 133, App
2	571.6	49.1	2243	5	US-09-397-206A-5474	Sequence 5474, Ap
c 3	185.6	16.0	265	5	US-09-397-206A-4445	Sequence 4445, Ap
c 4	185.6	15.9	289	5	US-09-397-206A-4472	Sequence 4472, Ap
5	183.6	15.8	553	5	US-09-397-206A-1697	Sequence 1697, Ap
6	156.2	13.4	681	7	US-10-488-619-1828	Sequence 1828, Ap
7	155.6	13.4	641	5	US-09-397-206A-3323	Sequence 3323, Ap
8	155	13.3	538	5	US-09-397-206A-3325	Sequence 3325, Ap
9	152.4	13.1	1345	6	US-10-501-035-7	Sequence 7, Appli
10	152.4	13.1	1345	9	US-60-584-405-119	Sequence 119, App
11	142	12.2	1691	7	US-10-804-491-24	Sequence 24, Appl
12	142	12.2	1923	7	US-10-100-683-944	Sequence 944, App
13	139	11.9	753	1	PCT-US04-23166-257	Sequence 257, Appl
14	139	11.9	2753	1	PCT-US04-14618-74	Sequence 74, Appl
15	136.4	11.7	2781	1	PCT-US04-23166-352	Sequence 352, App
16	135.8	11.7	3722	6	US-10-501-933-3396	Sequence 3396, Ap
17	134.8	11.6	440	5	US-09-397-206A-4433	Sequence 4433, Ap
18	134.8	11.6	1489	1	PCT-US04-23166-14	Sequence 14, Appl
19	134.8	11.6	2919	1	PCT-US04-23166-494	Sequence 494, Appl
20	132.6	11.4	3327	7	US-10-487-092-60	Sequence 60, Appl
21	129.8	11.2	1398	1	PCT-US03-27382-25	Sequence 25, Appl
22	129.8	11.2	1398	1	PCT-US03-04688A-124	Sequence 124, Appl
23	122.6	10.5	2383	1	PCT-US03-04688A-595	Sequence 595, App
c 24	122.2	10.5	558	6	US-10-501-933-2472	Sequence 2472, Ap

25	120.6	10.4	955	7	US-10-220-964-3856	Sequence 3856, Ap
26	112.4	9.7	546	5	US-09-969-034-1963	Sequence 1963, Ap
27	108.4	9.3	623	5	US-09-969-034-3013	Sequence 3013, Ap
c 28	108.2	9.3	1978	6	US-10-896-164-1408	Sequence 1408, Ap
29	108.2	9.3	1978	8	US-10-868-184-1408	Sequence 1408, Ap
30	108.2	9.3	1978	8	US-10-868-184A-1408	Sequence 1408, Ap
31	104.8	9.0	480	9	US-60-568-635-93	Sequence 93, Appl
32	99.6	8.6	1813	8	US-10-115-635-226	Sequence 226, Appl
33	99	8.5	663	1	PCT-US03-03403-47	Sequence 47, Appl
34	94.6	8.1	457	7	PCT-US03-04688A-126	Sequence 126, App
c 35	94.6	8.1	790	7	US-10-488-619-2358	Sequence 2358, Ap
36	94.4	8.1	452	6	US-10-425-115-71747	Sequence 71747, A
c 37	93.4	8.0	1768	1	PCT-US03-41761-53991	Sequence 53991, A
c 38	93.4	8.0	1768	1	PCT-US03-41761-53991	Sequence 53991, A
c 39	93.4	8.0	1768	1	PCT-US03-41766A-53991	Sequence 53991, A
c 40	89	7.6	591	7	US-10-834-268-4244	Sequence 4244, Ap
41	89	7.6	613	9	US-60-592-408-669	Sequence 669, App
42	83.8	7.2	342	5	US-09-796-692B-9309	Sequence 9309, App
43	83.8	7.2	342	5	US-10-764-324-9309	Sequence 9309, App
c 44	82.6	7.1	342	5	US-09-796-692B-306	Sequence 306, App
c 45	82.6	7.1	342	5	US-09-796-692B-458	Sequence 458, App

ALIGNMENTS

```
RESULT 1
US-10-115-635-133
; Sequence 133, Application US/10115635
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Zhou, Ping
;  APPLICANT: Goodrich, Ryle
;  APPLICANT: Liu, Chenghua
;  APPLICANT: Asundi, Vinod
;  APPLICANT: Ren, Feiyan
;  APPLICANT: Zhang, Jie
;  APPLICANT: Zhao, Qing A.
;  APPLICANT: Xue, Aidong J.
;  APPLICANT: Yang, Yonghong
;  APPLICANT: Wehrman, Tom
;  APPLICANT: Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION: Polypeptides
;  FILE REFERENCE: 797CON
;  CURRENT APPLICATION NUMBER: US/10/115,635
;  CURRENT FILING DATE: 2002-04-03
;  PRIOR APPLICATION NUMBER: 09/714,936
;  PRIOR FILING DATE: 2000-11-17
;  NUMBER OF SEQ ID NOS: 362
;  SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 133
;  LENGTH: 1884
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (157)..(864)
US-10-115-635-133

Query Match          51.8%;  Score 602.6;  DB 8;  Length 1884;
Best Local Similarity  75.3%;  Pred: No. 9.6e-126;
Matches  817; Conservative  0; Mismatches 234; Indels  34; Gaps  4;

Qy       1 TTTTTGTAAGAAAATCTGAGGAAAGATTCGGGATGAGCGCTTTATTC-AGGATGTTTTCC 59
             |||||||||||||||||||||||||||||||||  |||||||||||| |||||||||||
Db      25 TTCTTTTTTGTAAACATTACTGTAGGAGTTGTGATTAACTTTTTATTCTACTATGTATATG 84

Qy      60 TATGAAATAGCATTCATCTGTCGGGAGAGAAGGACTAAGGAAATCTGACATCGTTGGT 119
             |||||||||  ||||  ||  ||||||  |||||||||  |||||||
Db      85 TATGGAATAGTATTAATAAATGAACTAGGGAAGGATGTAATAAAATTAGACATCCTTCAT 144

Qy     120 CACTGGGACAGAATATGGATATGTCACTGGTTTCTCTCTCAATTAAATATGTGTGAAT 179
```

```
;  APPLICANT: Holtzman, Douglas A.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
;  TITLE OF INVENTION: HUMAN FETAL SPLEEN LIBRARY
;  FILE REFERENCE: MLN98-41pM
;  CURRENT APPLICATION NUMBER: US/09/397,206A
;  CURRENT FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: 60/101,133
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/107,254
;  PRIOR FILING DATE: 1998-11-05
;  PRIOR APPLICATION NUMBER: 60/126,842
;  PRIOR FILING DATE: 1999-03-30
;  NUMBER OF SEQ ID NOS: 6473
;  SOFTWARE: FastSEQ for Windows Version 3.0
;  SEQ ID NO 5474
;  LENGTH: 2243
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (1)...(2243)
;  OTHER INFORMATION: n = A,T,C or G
US-09-397-206A-5474

Query Match          49.1%;  Score 571.6;  DB 5;  Length 2243;
Best Local Similarity 77.3%;  Pred. No. 9.3e-119;
Matches  733;  Conservative   0;  Mismatches 184;  Indels  31;  Gaps   2;

Qy    137 GATATGTCACTGGTTTCTCTCTCAATAAATATGTGAATCAAAAGAGAAACAAGAAAACTTTT 196
              |||    ||||||  | |||||| | ||||||||| | || |||||||||||||||||
Db    152 GAAACAACATTGCTTTCTTTTCTTTTCAAATAAATAAATATGTGAATCAAAAGAGAAAAACTTTT 211

Qy    197 TTCAAACTAATGCATGGTCAGGGAAGAAGAAACAAGCATCAGAGGCCAAAATCAGCAGAGCG 256
              ||||||||||| ||| |||||||| |||||||||||||||||||||||||||| |||||
Db    212 TTCAGTTAATACATGGTTCAGGAAAAAGAAGAAGAAACAAGCAAAATCAGAGCT 271

Qy    257 AAAGAAAAAAGGGATAGACTAAGTCTTCCTACCAGAGGCCTGACTTCCATGAGAGACT 316
              |||||||||| |||||||| |||||||||||| ||||||||||| |||||||||||
Db    272 AAGGAAAAAGAAATAGAATAGACTAAGTCTTCTTGTGCGAGAAACCTGAGTTTCATGAGAGACACC 331

Qy    317 CAAGCCAGTAGATCTGCCCTTCGGCCAAAGAGTCTCTCCTCCTGCCAAAGAAACAAGAGTCAGTG 376
              |||||||||||| || ||||| |||||||||||||||||||||||||||||||||||||
Db    332 CGCTCCAGTAGATCTGGGCACTTGGCCAAAGAGAAACAAGAGTCTCCCCTGAAGAGGCAGTG 391

Qy    377 AAATGGGCTGAATCATTGACAAATGCTTCTCCACAGTTTTGTTACC 436
              ||||||||||||||||||||||||||| || || |||||||||||
Db    392 AAATGGGGTGAATCATTGACAAACTGCTTTCCCATAGAGATGACTAGGAGGCTTTTTACC 451

Qy    437 AGATTTCTTAAAACTGAATTCAGTGAGAGAACATTGATTTTGGGTCGCCCTGTGAAGAC 496
              |||||| |||||||||||||||||||||||||||||||||||||||  |||||||||
Db    452 AGATTTCTTAAAACTGAATTCAGTGAAGAGAAAAATATTGAATTTTGGATAGCCTGTGAAGAT 511

Qy    497 TTCAGAAATGCAAGGAACCTCAGAGGTAAAGGCAATCATCCTAAAAGCAAAGGCAATCATGAGAAA 556
              ||||||||||||||||||||||| ||||||| ||||||| |||||||||||||||||||
Db    512 TTCAGAAATGCAAGGGAACCTCAGAGGGTAAAGGCAATCATGAGAAA 571

Qy    557 TTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTTCATACTAAAGAAGTAATT 616
              | || ||| ||| |||||| |||||||||||||||||||||| |||||||||||||||
Db    572 TTTATACAGACTGATGCCCCCAAAGAGGTTAACCTTGATTTTCACACAACAAAGAGTCATT 631

Qy    617 GCTAAGAGCATCGGCCAGCCCACTCTCCACAGTTTTGATACGGCACAAAGCAAAGTGTAC 676
              ||||  |||||||||||||||||||| |||| ||||| ||||| || ||||||||||
Db    632 ACAAACAGCATCACTCACCTCACCTACCCTCCACAGTTTTGATGCTGCACAAAGCAGAGTGTAT 691

Qy    677 CAGCTCATGAAACATGACAGTTATAAACGCTTTTGAAATCTGAAATCTTTACATTTG 736
              ||||||||||||||||||||| || || ||||||||||||||||| ||||||||
Db    692 CAGCTCATGAAGGAAACAAGAACAGTTATACACGTTTCTGAAATCTGAAATCTTAGACTTG 751

Qy    737 ATAGAAGAAGACCTCAGACGCCAACAAACCTTAGGACGATCACGATCATTACTTAC 796
              |||||||||| ||  ||  ||||||||||| ||||  ||||||| ||||||||
Db    752 ATGGAAGGAAGACCTCAGAGACCAGAGACCAACAAATCTTAGGAGAGCATCACGCTCATTACCTGC 811

Qy    797 AATGAGATTTCCAAGATGTAAAGTAAAGTCAGATGTTGCCATTTGGTTATGATGGTAAAAGTCATTTGT 856
```

```
Db     145 T--TTAGAGAGAAGAATGAAACAACATTGCTTTCTTTTTCTCAAATAAATATGTGAAT 202
               ||  ||||  ||||||||||||||||||||| |||| ||||||||||||||||
Qy     180 CAAAAGAGAAAAACTTTTTCAAACTAATGCATGGGTCAGGGAAAAGAAGAAGAAACAAGCATCG 239
               ||||| |||||||||||| ||||||||||| |||||||||| ||||||||||||||
Db     203 CAAAAGAAAAAACTTTTTTCAGTTAATACATGGTTCAGGAAAAAGAAAAGAAACAAGCAAAG 262

Qy     240 AGGCCAAAATCAGACGCGAAAGAAAAAAGGAATAGACTAAGTCTTCCTACCAGAGGCCTG 299
               | || |||||||||| ||||||||||||  ||| || ||||||||| | ||||||||||
Db     263 AAGCCAAAATCAGAGCTGTAAGGAAAAAGAAATAGACTAAGTCTTCTTGTGCGAGAAACCTG 322

Qy     300 ACTTCCATGGAGAGACTCAAGCCAGTAGAATGTGCCCTTCGGCCAAAGAAACAAGAGTCT 359
               ||||||||| |||||||||||||||||| |||||||||||||||||||| ||||||||||
Db     323 AGTTTCATGAAGAGACACACCCGCTCCAGTAGATCTGGGGCACTTGGCCAAAGAAACAAGAGTCT 382

Qy     360 CTCCTGAAGAAGCAGTGAAATGGCTGAATCATTGACAAATGCTTCTCTCCATAGAGATG 419
               |||| ||||||||||| ||||  | |||||||||||||||||||||||| || | |||
Db     383 CCCCTGAAGAGGCAGTGAAATGGGGTGAATCATTGACAAACTGCTTTCCCATAGAGATG 442

Qy     420 GAGTGGATGCGCTTTACCAGATTCTTAAAAACTGAATTCAGTGAGAGAACATTGATTTT 479
               ||||||||| || ||||||||| |||||||||||||||||||||||||||||||||||
Db     443 GACTAGGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAAGAGAAAAATATTGAATTTT 502

Qy     480 GGGTCGCCCTGTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAA 539
               ||||  |||||||||||||||| |||||||||||||||||||||| ||||| |||||||||||
Db     503 GGATAGCCTGTGTGAAGAATTTCAGAAATGCAAGGGAACCTCAACAAATTCACCTTAAAGCAA 562

Qy     540 AGCAATCTATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTTC 599
               ||| |||| || ||| ||| || |||||||||||| |||||||||||||| ||||||||
Db     563 AAGCAATAATATGAGAAATTATACAGACTGATGCCCCCAAAGAGGTTAACCTTGATTTTC 622

Qy     600 ATACTAAAGAAGTAATTGCTAAGAGCATCGGCCAGCCCACTCTCCACAGTTTTGATACGG 659
               |||| ||||| |||| ||||  |||||||| || ||||||||| |||||||||| |||||
Db     623 ACACCAAAGAGTCATTACAACAGCATCACTCACCTCACCTACCCTCCACAGTTTTGATGCTG 682

Qy     660 CACAAAGCAGAGTGTACCAGCTCATGAAACATGACAGTTATAAACGCTTTTGAAATCTG 719
               ||||||||| ||||| ||||||||||| |||| || |||| || || |||||||||||
Db     683 CACAAAGCAGAGTGTATCAGCTCATGAAGGAAACAAGAACAGTTATACACGTTTCTGAAATCTG 742

Qy     720 AGACCTACTTACATTTGATAGAAGGAAGACCTCAGAGACCAACAAACCTTAGGACGACGAT 779
               | || || ||||||||||| ||||||||||||||||||||||||| |||| ||||||||
Db     743 ACATCTATTTGACTTGATGGAAGGAAGACCTCAGAGACCAGAGACCAACAAATCTTAGGAGAGCGAT 802

Qy     780 CACGATCATTACTTACACAATGATTCCAAGATGTAAAGTGAAAGTCAGATGTTGCCATTTGGTTAT 839
               |||||| || |||||||||||||||||||| ||||| ||| || |||||||||||||||||
Db     803 CACGCTCATTACCTGCCAATGAATTCCAAGATGTACCAAGATGTTGCCATTTGGTTAT 862

Qy     840 GAGTAAAGTCATTTGTCTTCTTTGAGTAGTGTATATCTAAAATATATACTAAT 899
               ||| ||| |||| ||||| ||||| |||||||| || ||  |||||||
Db     863 AAAGAAAATTGATTTTGCTCATTTTTATGACAAACTTATACATC------------ 906

Qy     900 ACTAATGTGTACTTCTAAAATATAGCTTGTGTGTATAGAAGAAGATGATTTCATTTT------- 954
               |||  |||| || ||||| || |||||||||| | |||||||| || ||||||
Db     907 --------TGCTTCTAACATATCGCATGTTTATGTTAAGATTTGGTCCCATCCTTTAAA 957

Qy     955 -TAAAATACACCACGAAATACATATTAAATGTAAGAACTTTTTATATTATTACTAAAATA 1013
                ||||||| || |||| ||| || || |||||  ||| | |||||||| |||||||||||
Db     958 CTGAAATATGTCATGTGAAAAATTATTTTAAAAATGTAAAAACAAAACTTTCTGCTAACAAA 1017

Qy    1014 ATTCATCATCTATCTTCCGAAAATATTTTATGAAAATCTATCTGATATTCTATTCTAATAA 1073
               |||| |||| || ||||||||  |||||||  || ||| |||  || |||||||| ||
Db    1018 ATACATACAGATATCTGCCAGTATTATTTTAAAAACAACCTTCTATTTGATGTCATTCCATTTA 1077

Qy    1074 AATTC 1078
               ||||
Db    1078 TAATC 1082
```

Db   812 ||||| |||||||||| |||| ||||||||| | ||||||||||| ||||||||| ||||||
           AATGAATTCCAAGAGATGTACAATCAGATGTTGCCATTTGGTTATATAAAGAAAAATTGATTTTG 871

Qy   857 CTTCTTTTGATAGTGTATGTGTATATCTAAAATATATACTAATACTAATGTGTACTTCTA 916
           |||||||| |||||
Db   872 CTCATTTTTATGACAAACTTATACATC------------------------TGCTTCTA 906

Qy   917 AAATATAGCTTGTGTATAAGAGAGATGATTTCATTT------TAAAATACACCATGCA 970
           || ||||| |||||| ||| || |||||
Db   907 ACATATCGCATGTTATGCTTAAGAGATTGGTCCCATCCTTAAACTGAAAATAGTCATGTG 966

Qy   971 AATACATATTAAAATGTAAGAACTTTTTATATATTATACTAAAATAATCATCATCTATCTTC 1030
           | ||| ||| |||| | ||| ||||| || || ||||
Db   967 AAATTATTTTAAAAATGTAAAAACAAAACTTTCTGCTAACAAAATACATACAGTATCTGC 1026

Qy   1031 CGAAAATATTTATGAAAATCTATCTGATATTCGATTCATTCCTAATCAATAAAATC 1078
           | | |||| ||||| || | ||| |||| |||| || ||||||
Db   1027 CAGTATATTCTGTAAAACCTTCTATTTGATGTCATTCCATTTATAATC 1074

RESULT 3
US-09-397-206A-4445/c
; Sequence 4445, Application US/09397206A
; GENERAL INFORMATION:
;   APPLICANT: Holtzman, Douglas A.
;   APPLICANT: Gearing, David P.
;   TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
;   TITLE OF INVENTION: HUMAN FETAL SPLEEN LIBRARY
;   FILE REFERENCE: MLN98-41pM
;   CURRENT APPLICATION NUMBER: US/09/397,206A
;   CURRENT FILING DATE: 1999-09-17
;   PRIOR APPLICATION NUMBER: 60/101,133
;   PRIOR FILING DATE: 1998-09-18
;   PRIOR APPLICATION NUMBER: 60/107,254
;   PRIOR FILING DATE: 1998-11-05
;   PRIOR APPLICATION NUMBER: 60/126,842
;   PRIOR FILING DATE: 1999-03-30
;   NUMBER OF SEQ ID NOS: 6473
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4445
;   LENGTH: 265
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-397-206A-4445

Query Match   16.0%; DB 5; Length 265;
Best Local Similarity 86.5%; Score 185.8; Pred. No. 3.2e-32;
Matches 205; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy   628 CGCCCAGCCCACTCTCCACAGTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGA 687
           |||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||
Db   257 CGCTCAACCTACCCTCCACAGTTTGATCTGCGCAAAGCAGAGTGTATCAGCTCATGGA 198

Qy   688 ACATGACAGTTATAAACGCTTTTGAAATCTGAGAATCTGAGATAGAAGGAAG 747
           ||| |||||||||| ||| ||||||| |||||| || ||||||||||||||
Db   197 ACAAGACAGTTATACACGTTTTCTGAAATCTGAAATCTCATTATTGATGGAAGGAAG 138

Qy   748 ACCTCAGAGACCAACAAACCTTAGGAGACGATCACGATCATTCACTTACTTACAATGATTTCA 807
           ||||||||||||||||||||||||||||||| ||||||| || |||||| | ||
Db   137 ACCTCAGAGACCAACAAAATCTTAGGAGACGATCACGCTCATTTACCTGCAATGAATTCCA 78

Qy   808 AGATGTAAAGTCAGATGTTGCCATTGGTTATGAGTAAAAGTCATTGTCTTCTTTT 864
           |||||||| |||| ||||||||||||||||||| ||||| ||||| | |||||
Db   77 AGATGTACAATCAGATGTTGCCATTTGGTTATAAAGAAAATTGATTTTGCTTCATTT 21

RESULT 4
US-09-397-206A-4472/c
; Sequence 4472, Application US/09397206A
; GENERAL INFORMATION:
;   APPLICANT: Holtzman, Douglas A.
;   APPLICANT: Gearing, David P.
;   TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A

— — —

;   TITLE OF INVENTION: HUMAN FETAL SPLEEN LIBRARY
;   FILE REFERENCE: MLN98-41pM
;   CURRENT APPLICATION NUMBER: US/09/397,206A
;   CURRENT FILING DATE: 1999-09-17
;   PRIOR APPLICATION NUMBER: 60/101,133
;   PRIOR FILING DATE: 1998-09-18
;   PRIOR APPLICATION NUMBER: 60/107,254
;   PRIOR FILING DATE: 1998-11-05
;   PRIOR APPLICATION NUMBER: 60/126,842
;   PRIOR FILING DATE: 1999-03-30
;   NUMBER OF SEQ ID NOS: 6473
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4472
;   LENGTH: 289
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-397-206A-4472

Query Match   15.9%; DB 5; Length 289;
Best Local Similarity 85.8%; Score 185.6; Pred. No. 3.6e-32;
Matches 206; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy   628 CGCCCAGCCCACTCTCCACAGTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGA 687
           |||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||
Db   252 CGCTCAACCTACCCTCCACAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGA 193

Qy   688 ACATGACAGTTATAAACGCTTTTGAAATCTGAGAATCTGAGATAGAAGGAAG 747
           ||| |||||||||| ||| ||||||| |||||| || ||||||||||||||
Db   192 ACAAGACAGTTATACACGTTTTCTGAAATCTGAAATCTGACATCCATTAGACTTGATGGAAGGAAG 133

Qy   748 ACCTCAGAGACCAACAAACCTTAGGAGACGATCACGATCATTCACTTACTTACAATGATTTCCA 807
           ||||||||||||||||||||||||||||||| ||||||| || |||||| | ||
Db   132 ACCTCAGAGACCAACAAAATCTTAGGAGACGATCACGGATCATTTACCTGCAATGAATTCCA 73

Qy   808 AGATGTAAAGTCAGATGTTGCCATTGGTTATGAGTAAAAGTCATTGTCTTCTTTTGAT 867
           |||||||| |||| ||||||||||||||||||| ||||| ||||| | |||||| |
Db   72 AGATGTACAATCAGATGTTGCCATTTGGTTATAAAGAAAATTGATTTTGCTCATTTTAT 13

RESULT 5
US-09-397-206A-1697
; Sequence 1697, Application US/09397206A
; GENERAL INFORMATION:
;   APPLICANT: Gearing, David P.
;   APPLICANT: Holtzman, Douglas A.
;   TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
;   TITLE OF INVENTION: HUMAN FETAL SPLEEN LIBRARY
;   FILE REFERENCE: MLN98-41pM
;   CURRENT APPLICATION NUMBER: US/09/397,206A
;   CURRENT FILING DATE: 1999-09-17
;   PRIOR APPLICATION NUMBER: 60/101,133
;   PRIOR FILING DATE: 1998-09-18
;   PRIOR APPLICATION NUMBER: 60/107,254
;   PRIOR FILING DATE: 1998-11-05
;   PRIOR APPLICATION NUMBER: 60/126,842
;   PRIOR FILING DATE: 1999-03-30
;   NUMBER OF SEQ ID NOS: 6473
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1697
;   LENGTH: 553
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(553)
;   OTHER INFORMATION: n = A,T,C or G
US-09-397-206A-1697

Query Match   15.8%; Score 183.6; DB 5; Length 553;

```
Best Local Similarity  68.0%;  Pred.No. 1.2e-31;
Matches  300;  Conservative  0;  Mismatches  110;  Indels  31;  Gaps  2;

Qy   644  CACAGTTTGATGGGCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAA   703
              ||||||||||| |||||||||| |||||||||| |||||||||||||||||||
Db     9  CCCAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCATGGAACAAGACAGTTATACA    68

Qy   704  CGCTTTTGAAATCTGAGACCTACTTACATTTGATAGAAGGAGACCTCAGAGACCAACA   763
              || ||| ||||||||| |||  ||||||||||||||||||||||||||||||||
Db    69  CGTTTTCTGAAATCTGACACTCTCATTTAGACTTAGCTTCTACTACTAGAGACCAACA   128

Qy   764  AACCTTAGAGAGACGATCACGATCCATTCATTTACCTGCAATGAATGTGTACAATCAGAT   823
              |||||||||||||||          |||||||||||||||||||||||||||||
Db   129  AATCTTAGAGAGACGATCCATGGATGAATTCCAAGAGTGTACAATCAGAT   188

Qy   824  GTTGCCATTGGTATGAGTAAAAGTCATTTGTCTTTTTGATAGTGTATGTGTATC   883
              ||| ||||||   ||||||||
Db   189  GTTGCCATTGGTTATAAAGAAGAAAATTGATTTTGCTCATTTTATGACAAACTTATACAATC   248

Qy   884  TAAAATATATACTAATACTAATGTGTACTTCAAAAATATAGCTTGTGTATAAGAAGAGAT   943
                                        TGCTTCTAACATATCGCATGTTTATGTTAAGATTT   283
Db   249  --------------------------

Qy   944  GATTTCATTTT------TAAAATACACCATGCAAATACATATTAAATGTAAGAACTTTTT   997
Db   284  GGTCCCATCCTTTAAACTGAAATATGTCATGTGAAATTATTTTAAAAATGTNAAAACAAA   343

Qy   998  ATATTATACTAAAATAATTCATCATCTATCTTCCGAAATATTTTATGAAAATCTATCTGA  1057
Db   344  ACTTTCTGCTAACAAAATACATACAGTATCTGCCAGTAGCTATAGTTCTGTAAAACCTTCTATTT   403

Qy  1058  TATTCTATTCTAATAAAATTC  1078
              ||||||||||||
Db   404  GATGTCATTCCATTTATAATC   424
```

```
RESULT 6
US-10-488-619-1828
; Sequence 1828, Application US/10488619
; GENERAL INFORMATION:
;   APPLICANT: Greenlee, Winner and Sullivan, P.C.
;   TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations
;   TITLE OF INVENTION: Physiological Conditions, And Genotyping Arrays
;   FILE REFERENCE: 98-01 WO
;   CURRENT APPLICATION NUMBER: US/10/488,619
;   CURRENT FILING DATE: 2004-03-01
;   NUMBER OF SEQ ID NOS: 3040
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 1828
;   LENGTH: 681
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-488-619-1828

Query Match  13.4%;  Score 156.2;  DB 7;  Length 681;
Best Local Similarity  64.0%;  Pred.No. 1.8e-25;
Matches  236;  Conservative  0;  Mismatches  133;  Indels  0;  Gaps  0;

Qy   359  TCTCCTGAAGAAGCAGTGAAATGGCTGAATCATTTGACAAATTGCCTCTCATCAGAGAT   418
              ||||||||||||||||||||||||||||||||    |||||||||||
Db   220  TCTCCTGAGGAAGCGCAGCTCTGGGCAGAAGCATTTGATGAACTGCTGGCCAGTAAATAT   279

Qy   419  GGAGTGGATGCTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT   478
              |||  ||||||||||||      |||  |||  ||||||
Db   280  GGGCTGGCTGCATTCAGGGCGTTTCAGGGCCGTTCTGAGTCCGAGTTCTGTGAAGAAAACATTGAATTC   339

Qy   479  TGGGTCGCCTGTGTGAAGACTTCAAGAAATGCAAGAACCTCAACAAATCATCCTAAAAGCA   538
              |||||||| |||||| ||||||
Db   340  TGGTTGGCTTGTGAAGACTTCAAGAAAAACCAAAAAACTGTCCTCAAAAGCA   399

Qy   539  AAGGCAATCTATGAGAAATTCATTCAGATGATGCCCCCAAAGAGGTTAACATTGATTTT   598
              ||||||||||||||  |||||||| ||||||||||||
Db   400  AGGAAAAATCTATACCGACCTTCATAGAGAAGAGAAGCCTCCAAAGAGATAAACATAGACTTC   459
```

```
RESULT 7
US-09-397-206A-3323
; Sequence 3323, Application US/09397206A
; GENERAL INFORMATION:
;   APPLICANT: Gearing, David P.
;   APPLICANT: Holtzman, Douglas A.
;   TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
;   TITLE OF INVENTION: HUMAN FETAL SPLEEN LIBRARY
;   FILE REFERENCE: MLN98-41pM
;   CURRENT APPLICATION NUMBER: US/09/397,206A
;   CURRENT FILING DATE: 1999-09-17
;   PRIOR APPLICATION NUMBER: 60/101,133
;   PRIOR FILING DATE: 1998-09-18
;   PRIOR APPLICATION NUMBER: 60/107,254
;   PRIOR FILING DATE: 1998-11-05
;   PRIOR APPLICATION NUMBER: 60/126,842
;   PRIOR FILING DATE: 1999-03-30
;   NUMBER OF SEQ ID NOS: 6473
;   SOFTWARE: FastSEQ for Windows Version 3.0
;   SEQ ID NO 3323
;   LENGTH: 641
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(641)
;   OTHER INFORMATION: n = A,T,C or G
US-09-397-206A-3323

Query Match  13.4%;  Score 155.6;  DB 5;  Length 641;
Best Local Similarity  82.7%;  Pred. No. 2.4e-25;
Matches  201;  Conservative  0;  Mismatches  40;  Indels  2;  Gaps  2;

Qy   644  CACAGTTTGATACGGCACAAAGCA-GAGTGTACCA-GCTCATGGAACATGACAGTTATA   701
              ||||||||| ||||||||||||  |||||||||  |||||||||||||  |||||||
Db     7  CNCAGTTTTGATGCTGCACAAAGCACGAGTGTATCACGCGTCATGGAACAACAAGAGTTATA    66

Qy   702  AACGGCTTTTGAAATCTGAGACCTACTTACATTTGATAGAAGGAAGACCTCAGAGACCAA   761
              ||  |||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db    67  CACGTTTTCTGAAATCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAAGACCTCAGAGACCAA   126

Qy   762  CAAACCTTAGAGAGACGATGATCACGATCCATTTACTTACTACCATCAATCAG   821
              ||||||||||||||||| |||||  |||||| ||||||  ||||||||||||
Db   127  CAAATCTTAGAGAGACGATCACGCCTCATTTACCTGCAATGAATGTACAATCAG   186

Qy   822  ATGTTGCCATTGGTTATGAGTAAAAGTCATTTGTTCTTCTTTGATAGTGTATGTGTATA   881
              ||||||||||||||||  ||||||  ||||||||
Db   187  ATGTTGCCATTTGGTTATAAAGAAAAATTGGATTTTGCTCATTTTATGACAAACTTATACA   246

Qy   882  TCT   884
              |||
Db   247  TCT   249
```

```
RESULT 8
US-09-397-206A-3325
; Sequence 3325, Application US/09397206A
; GENERAL INFORMATION:
;   APPLICANT: Gearing, David P.
```

```
Qy   599  CATACTAAAGAAGTAATTGCTAAGAGCATCGCCCACTCTCCACAGTTTGATACG   658
              |||||||||||||||||||||  ||  ||
Db   460  CAAACGAAATCTCTGATTGCCCAAATATCCAAGAGGCTACAAGTGGCTGCTTCCACCACA   519

Qy   659  GCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTGAAATCT   718
              |||||||||||||||||||||||||||||||||||||||||||||||
Db   520  GCTCAGAAGAGGGTGTACAGTTTGATGGGAGAACAATTCTTATCCTCGGTTCTTGGAGTCC   579

Qy   719  GAGACCTAC   727
              |||||||||
Db   580  GAATTCTAC   588
```

; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN FETAL SPLEEN LIBRARY
; FILE REFERENCE: MLN98-41pM
; CURRENT APPLICATION NUMBER: US/09/397,206A
; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/101,133
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/107,254
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/126,842
; PRIOR FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 6473
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3325
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(538)
; OTHER INFORMATION: n = A,T,C or G
US-09-397-206A-3325

Query Match   13.3%; Score 155; DB 5; Length 538;
Best Local Similarity   83.2%; Pred. No. 3.2e-25;
Matches   188; Conservative   0; Mismatches   36; Indels   2; Gaps   1;

Qy   644 CACAGTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAAA 703
               |||||||||| ||||| |||| |||| | || ||| ||||||||||| ||| ||||
Db   4 CCCAGTTTGATGCTGCACAAAGCAGAGTGTATCAGCTCAGCAGACAGTTATAC   63

Qy   704 --CGCTTTTTGAAATCTGAGACCTACTTACATTTGATAGAGGAAGGAACCTCAGAGACCAA 761
             ||||||||||||||| ||||||| || |||||| ||||||||||||| |||||| |||
Db   64 ACCGTTTTCTGAAATCTGACATCTATTTAGACTTGATGGAAGGAAGGACCTCAGAGACCAA 123

Qy   762 CAAACCTTAGGAGAGACGAATCACGATCACCATCATTTACTTACAATGATTTCCAAGATGTCAAAGTCAG 821
               ||||||||||||||| ||||| |||| || || |||||| ||||||||||||| |||||
Db   124 CAAATCTTAGGAGAGACGAATCACCGTCATTTACCTGCAATGAATTCCAATGATGTACAATCAG 183

Qy   822 ATGTTGCCATTTGGTTATGAGTAAAAGTCATTTGTCTTCTTTTGAT 867
               |||||||||||||||||| |||| ||||||||| ||| ||||
Db   184 ATGTTGCCATTTGGTTATTGGTTATAAAGAAAAATTGATTTTGCTCATTTTAT 229

RESULT 9
US-10-501-035-7
; Sequence 7, Application US/10501035
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE
; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS
; FILE REFERENCE: D0185 PCT
; CURRENT APPLICATION NUMBER: US/10/501,035
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: US 60/350,061
; PRIOR FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 1345
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-501-035-7

Query Match   13.1%; Score 152.4; DB 6; Length 1345;
Best Local Similarity   62.7%; Pred. No. 1.5e-24;
Matches   237; Conservative   0; Mismatches   141; Indels   0; Gaps   0;

Qy   359 TCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGAT 418
               |||||||||||||| || ||||||| ||||||||||| || ||||| |||||||
Db   249 TCTCCTGAGGAAGCACAGCTGTGGTCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy   419 GGAGTGGATGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 478
               |||||||||| | |||||||||||||||||||||||||||||||||||||||
Db   309 GGTCTTGCTGCATTCAGGGCTTTTTAAGTCGGAATTCTGTGAGAAACATATTGAATTC 368

Qy   479 TGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGAAACCTCAACAAATCATCCTAAAAGCA 538
               ||| ||||||||||||||||||||| |||||||| ||||| |||| |||||||||||
Db   369 TGGCTGGCCTGTGAAGACTTCAAAAAACCAAATCACCCCAAAGAGTGTCTCAAAAGCA 428

Qy   539 AAGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTT 598
               |||||||||||| ||| |||| ||||||||| ||| ||||| |||||||||||||
Db   429 AGGAAAATATATACTGACTTCATAGAAAAGGAAGCTCAAAAGAGATAAACATAGATTTT 488

Qy   599 CATACTAAAGAAGTAATTGCTAAGAGCATGCCCAGCCCACTCTCCACAGTTTGATACG 658
               || |||||| |||||||| |||||||||||| ||| ||||||||||||||||||||||
Db   489 CAAACCAAAACTCTGATTGCCCAGAATATACAAGGAGCTACAAGTGGCTGCTTTACAACT 548

Qy   659 GCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAAACGCTTTTTGAAATCT 718
               |||||||| | || |||| || | ||||||||||| | ||||||| |||||||||||
Db   549 GCCCAGAAAAGGTATACAGCTTGATGGAGAACAACTCTTATCCTCGTTTTCTTGGAGTCA 608

Qy   719 GAGACCTACTTACATTTG 736
               || ||| ||| |||
Db   609 GAATTCTACCAGGACTTG 626

RESULT 10
US-60-584-405-119
; Sequence 119, Application US/60584405
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
; TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE
; TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 10001 PSP
; CURRENT APPLICATION NUMBER: US/60/584,405
; CURRENT FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 527
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 119
; LENGTH: 1345
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-584-405-119

Query Match   13.1%; Score 152.4; DB 9; Length 1345;
Best Local Similarity   62.7%; Pred. No. 1.5e-24;
Matches   237; Conservative   0; Mismatches   141; Indels   0; Gaps   0;

Qy   359 TCTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGAT 418
               |||||||||||||| || ||||||| ||||||||||| || ||||| |||||||
Db   249 TCTCCTGAGGAAGCACAGCTGTGGTCAGAAGCATTTGACGAGCTGCTAGCCAGCAAATAT 308

Qy   419 GGAGTGGATGCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 478
               |||||||||| | |||||||||||||||||||||||||||||||||||||||
Db   309 GGTCTTGCTGCATTCAGGGCTTTTTAAGTCGGAATTCTGTGAGAAACATATTGAATTC 368

Qy   479 TGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGAAACCTCAACAAATCATCCTAAAAGCA 538
               ||| ||||||||||||||||||||| |||||||| ||||| |||| |||||||||||
Db   369 TGGCTGGCCTGTGAAGACTTCAAAAAACCAAATCACCCCAAAGAGTGTCTCAAAAGCA 428

Qy   539 AAGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTT 598
               |||||||||||| ||| |||| ||||||||| ||| ||||| |||||||||||||
Db   429 AGGAAAATATATACTGACTTCATAGAAAAGGAAGCTCAAAAGAGATAAACATAGATTTT 488

Qy   599 CATACTAAAGAAGTAATTGCTAAGAGCATGCCCAGCCCACTCTCCACAGTTTGATACG 658
               || |||||| |||||||| |||||||||||| ||| ||||||||||||||||||||||
Db   489 CAAACCAAAACTCTGATTGCCCAGAATATACAAGGAGCTACAAGTGGCTGCTTTACAACT 548

Qy   659 GCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAAACGCTTTTTGAAATCT 718
               |||||||| | || |||| || | ||||||||||| | ||||||| |||||||||||
Db   549 GCCCAGAAAAGGTATACAGCTTGATGGAGAACAACTCTTATCCTCGTTTTCTTGGAGTCA 608

```
Qy      719 GAGACCTACTTACATTTG 736
            ||||||||||||
Db      609 GAATTCTACCAGGACTTG 626

RESULT 11
US-10-804-491-24
; Sequence 24, Application US/10804491
; GENERAL INFORMATION:
;   APPLICANT: Cismowski, Mary
;   APPLICANT: Duzic, Emir
;   TITLE OF INVENTION: AGS Proteins and Nucleic Acid Molecules and Uses Therefor
;   FILE REFERENCE: 60388-A-PCT-US
;   CURRENT APPLICATION NUMBER: US/10/804,491
;   CURRENT FILING DATE: 2004-03-19
;   PRIOR APPLICATION NUMBER: US/09/709,103
;   PRIOR FILING DATE: 2000-11-08
;   NUMBER OF SEQ ID NOS: 73
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
;   LENGTH: 1691
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (45)..(587)
;   OTHER INFORMATION:
US-10-804-491-24

Query Match       12.2%;    Score 142;   DB 7;   Length 1691;
Best Local Similarity 59.8%;   Pred. No. 3.5e-22;
Matches  238;   Conservative   0;   Mismatches 160;   Indels   0;   Gaps   0;

Qy      341 GCCAAAGAAACAAGAGTCTCTCCTGAGAAGCAGCAGTGAAATGGCTGAATCATTGACAAA 400
            ||||||                   ||| |||||||||  |||  | ||||||||||||||
Db      186 GCCAAGACCCAGAAAACCTCGCTGGACGAGCCCTGCGTGCCTGAATTCCTGGACAAA 245

Qy      401 TTGCTCTCTCATAGAGATGGAGTGGATGCTTTTACCAGAGTTCTTAAAACTGAATTCAGT 460
            ||||||||||| ||| || ||| |||   |  |   |  || |||||||  ||||||||||
Db      246 CTCCTGCAGAACAACTATGGACTTGCCAGTTTCAAAAGTTTCCTGAAGTCTGAATTCAGT 305

Qy      461 GAGGAGAACATTGAATTTGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAA 520
            |||||| ||| |||||   | ||||  |  |  |  ||||   |  ||  ||| || |||
Db      306 GAGGAAAACCTTGAGTTCTGGATTGCCTGTGAGGATTGCCATGGAACAATCAAGTCCCTGCC 365

Qy      521 CAAATCATCCTAAAAGCAAGGCAATCTATGAGAAATTCATTCAGATGATGCGCCCCAAA 580
            ||||||||||| ||| ||||   ||| | |||  |  ||||   | ||||  |||||||
Db      366 AAGATGGCTGAGAAGGCAAAGCAAATTCATTGATGAGAAGAATCAATCAAACGGAGGCTCCTAAA 425

Qy      581 GAGGTAACATTGATTTCATACTAAAGAAGTAATTGCTAGAGAGCATCGCCCAGCCCACT 640
            ||||||  |   ||| |||  ||  | ||||||| || || |   || ||||||||||||
Db      426 GAGGTGAATATTGACCACTTCACTAAGGACACATCACAATGAGAAACCTGGTGGAACCTTCC 485

Qy      641 CTCCACAGTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTAT 700
            ||| |||||||||| | |   ||| ||||| |||||||||||| |||  ||||||||||
Db      486 CTGAGCAGCTTTGACATGGCCCAGAAAAGAATCCATGCCCTGATGGAAAAGGATTCTCTG 545

Qy      701 AAACGCTTTTTGAAATCTGAGACCTACTTACATTTGAT 738
            ||| ||||| || | ||||| || |||| || |||||
Db      546 CCTCGCTTTGTGTGCGCTCTGAGTTTTATCAGGAGTTAAT 583

RESULT 12
US-10-100-683-944
; Sequence 944, Application US/10100683
; GENERAL INFORMATION:
;   APPLICANT: Rosen, et al.
;   TITLE OF INVENTION: Human Secreted Proteins
;   FILE REFERENCE: PS900
;   CURRENT APPLICATION NUMBER: US/10/100,683
;   CURRENT FILING DATE: 2002-03-19
;   PRIOR APPLICATION NUMBER: US 60/040,162
;   PRIOR FILING DATE: 1997-03-07
;   PRIOR APPLICATION NUMBER: US 60/043,576
;   PRIOR FILING DATE: 1997-04-11
;   PRIOR APPLICATION NUMBER: US 60/047,601
;   PRIOR FILING DATE: 1997-05-23
;   PRIOR APPLICATION NUMBER: US 60/056,845
;   PRIOR FILING DATE: 1997-08-22
;   PRIOR APPLICATION NUMBER: US 60/043,580
;   PRIOR FILING DATE: 1997-04-11
;   PRIOR APPLICATION NUMBER: US 60/047,599
;   PRIOR FILING DATE: 1997-05-23
;   PRIOR APPLICATION NUMBER: US 60/056,664
;   PRIOR FILING DATE: 1997-08-22
;   PRIOR APPLICATION NUMBER: US 60/043,314
;   PRIOR FILING DATE: 1997-04-11
;   PRIOR APPLICATION NUMBER: US 60/047,632
;   PRIOR FILING DATE: 1997-05-23
;   PRIOR APPLICATION NUMBER: US 60/056,892
;   PRIOR FILING DATE: 1997-08-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13468
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 944
;   LENGTH: 1923
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (144)..(144)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1910)..(1910)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1912)..(1912)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-100-683-944

Query Match       12.2%;    Score 142;   DB 7;   Length 1923;
Best Local Similarity 59.8%;   Pred. No. 3.6e-22;
Matches  238;   Conservative   0;   Mismatches 160;   Indels   0;   Gaps   0;

Qy      341 GCCAAAGAAACAAGAGTCTCTCCTGAGAAGCAGCAGTGAAATGGCTGAATCATTGACAAA 400
            ||||||                   ||| |||||||||  |||  | ||||||||||||||
Db      406 GCCAAGACCCAGAAAACCTCGCTGGACGAGCCCTGCGTGCCTGAATTCCTGGACAAA 465

Qy      401 TTGCTCTCTCATAGAGATGGAGTGGATGCTTTTACCAGAGTTCTTAAAACTGAATTCAGT 460
            ||||||||||| ||| || ||| |||   |  |   |  || |||||||  ||||||||||
Db      466 CTCCTGCAGAACAACTATGGACTTGCCAGTTTCAAAAGTTTCCTGAAGTCTGAATTCAGT 525

Qy      461 GAGGAGAACATTGAATTTGGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAA 520
            |||||| ||| |||||   | ||||  |  |  |  ||||   |  ||  ||| || |||
Db      526 GAGGAAAACCTTGAGTTCTGGATTGCCTGTGAGGATTGCCATGGAACAATCAAGTCCCTGCC 585

Qy      521 CAAATCATCCTAAAAGCAAGGCAATCTATGAGAAATTCATTCAGATGATGCGCCCCAAA 580
            ||||||||||| ||| ||||   ||| | |||  |  ||||   | ||||  |||||||
Db      586 AAGATGGCTGAGAAGGCAAAGCAAATTCATTGATGAGAAGAATCAATCAAACGGAGGCTCCTAAA 645

Qy      581 GAGGTAACATTGATTTCATACTAAAGAAGTAATTGCTAGAGAGCATCGCCCAGCCCACT 640
            ||||||  |   ||| |||  ||  | ||||||| || || |   || ||||||||||||
Db      646 GAGGTGAATATTGACCACTTCACTAAGGACACATCACAATGAGAAACCTGGTGGAACCTTCC 705

Qy      641 CTCCACAGTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTAT 700
            ||| |||||||||| | |   ||| ||||| |||||||||||| |||  ||||||||||
Db      706 CTGAGCAGCTTTGACATGGCCCAGAAAAGAATCCATGCCCTGATGGAAAAGGATTCTCTG 765

Qy      701 AAACGCTTTTTGAAATCTGAGACCTACTTACATTTGAT 738
            ||| ||||| || | ||||| || |||| || |||||
Db      766 CCTCGCTTTGTGTGCGCTCTGAGTTTTATCAGGAGTTAAT 803
```

RESULT 13
PCT-US04-23166-257
; Sequence 257, Application PC/TUS0423166
; GENERAL INFORMATION:
; APPLICANT: Purdue Pharma L.P.
; APPLICANT: Jiefei Tong
; APPLICANT: Gang Jin
; APPLICANT: Rui-Ru Ji
; APPLICANT: Yixun Xu
; APPLICANT: Lillian W. Chiang
; APPLICANT: Daniel J. Lavery
; TITLE OF INVENTION: GENES ASSOCIATED WITH RESPONSES TO NEUROPATHIC PAIN
; FILE REFERENCE: 02755/200M584-WOO
; CURRENT APPLICATION NUMBER: PCT/US04/23166
; CURRENT FILING DATE: 2004-07-26
; PRIOR APPLICATION NUMBER: 60/485,101
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 868
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 257
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-23166-257

Query Match   11.9%;   Score 139; DB 1; Length 753;
Best Local Similarity 56.6%; Pred. No. 1.4e-21;
Matches 278; Conservative 0; Mismatches 210; Indels 3; Gaps 1;

```
Qy  251 AGAGCGAAAGAAAAAGGAATAGACTAAGTCTTCCTACAGAGGCCTGACTTCCATGGA  310
            ||  |||||  |||    |  |||||   ||  ||||||   || ||
Db  140 AGTGCAAAAGATATGAAACATCGGCTAGGTTTCTGCTGCAAAAATCTGATTCCTGTGAA  199

Qy  311 GAGA---CTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCTCTCCTGAA  367
Db  200 CACAATTCTTCCCACACACAAGAAGGACAAAGTGGTTATTTGCCAGAGAGTGAGCCAAGAG  259

Qy  368 GAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGATGGAGTGGAT  427
            ||||  ||||||||||||||||||
Db  260 GAAGTCAAGAAATGGGCTGAATCACTGAAAAACCTGATTAGTCATGAATGTGGGCTGCA  319

Qy  428 GCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTGGGTCGCC  487
            |||| |||  ||||||||  ||||||
Db  320 GCTTTCAAAGCTTTCTTGAAGTCTGAAGTCTGAATATAGTGAGGAGAATATTGACTTCTCGGATCAGC  379

Qy  488 TGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAAGGCAATC  547
            |||||||  ||||||||  ||
Db  380 TGTGAAGAGTACAAGAAAATCAAATCACCATCTAAACTAAGTCCCAAGGCCAAAAAGATC  439

Qy  548 TATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTTCATACTAAA  607
            ||| ||||||||||  ||||
Db  440 TATAATGAATTCATCTCAGTCCAGGCAACCAAAGAGGTGAACCTGGATTCTTGATACGGCACAAAGC  667

Qy  608 GAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTGATACGGCCACAAAGC  667
            |||| ||||  ||||
Db  500 GAAGAGACAAGCCGGAACATGCTACCACGGCCTACCACCGCTTGATGAGGGCCAGAAG  559

Qy  668 AGAGTGTACCAGCTCATGGAACATGGACACTGTGAAATCTGAGACCTAC  727
            |||||||  ||||
Db  560 AAGATTTTCAACCTGATGGAGAAGGATTCCTACCGGCGCTTCCCAAGTCTCGATTCTAT  619

Qy  728 TTACATTTGAT  738
            |||||||
Db  620 CTTGATTTGGT  630
```

RESULT 14
PCT-US04-14618-74
; Sequence 74, Application PC/TUS0414618
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Olsen, Nancy J
; APPLICANT: Aune, Thomas M

; TITLE OF INVENTION: A Gene Equation to Diagnose Rheumatoid Arthritis
; FILE REFERENCE: 1242/64 PCT
; CURRENT APPLICATION NUMBER: PCT/US04/14618
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: US 60/468,901
; PRIOR FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 74
; LENGTH: 2753
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-14618-74

Query Match   11.9%;   Score 139; DB 1; Length 2753;
Best Local Similarity 56.6%; Pred. No. 1.9e-21;
Matches 278; Conservative 0; Mismatches 210; Indels 3; Gaps 1;

```
Qy  251 AGAGCGAAAGAAAAAGGAATAGACTAAGTCTTCCTCCACAGAGGCCTGACTTCCATGGA  310
            ||  |||||  |||    |  |||||   ||  ||||||   || ||
Db  140 AGTGCAAAAGATATGAAACATCGGCTAGGTTTCTGCTGCAAAAATCTGATTCCTGTGAA  199

Qy  311 GAGA---CTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCTCTCCTGAA  367
Db  200 CACAATTCTTCCCACACACAAGAAGGACAAAGTGGTTATTTGCCAGAGAGTGAGCCAAGAG  259

Qy  368 GAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGATGGAGTGGAT  427
            ||||  ||||||||||||||||||
Db  260 GAAGTCAAGAAATGGGCTGAATCACTGAAAAACCTGAAAAACCTGATTAGTCATGAATGTGGGCTGCA  319

Qy  428 GCTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTGGGTCGCC  487
            |||| |||  ||||||||  ||||||
Db  320 GCTTTCAAAGCTTTCTTGAAGTCTGAAGTCTGAATATAGTGAGGAGAATATTGACTTCTCGGATCAGC  379

Qy  488 TGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAAGGCAATC  547
            |||||||  ||||||||  ||
Db  380 TGTGAAGAGTACAAGAAAATCAAATCACCATCTAAACTAAGTCCCAAGGCCAAAAAGATC  439

Qy  548 TATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTTCATACTAAA  607
            ||| ||||||||||  ||||
Db  440 TATAATGAATTCATCTCAGTCCAGGCAACCAAAGAGGTGAACCTGGATTCTTGATACGG  499

Qy  608 GAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTGATACGGCCACAAAGC  667
            |||| ||||  ||||
Db  500 GAAGAGACAAGCCGGAACATGCTACCACGGCCTACCACCGCTTGATGAGGGCCAGAAG  559

Qy  668 AGAGTGTACCAGCTCATGGAACATGGACACTGTGAAATCTGAGACCTAC  727
            |||||||  ||||
Db  560 AAGATTTTCAACCTGATGGAGAAGGATTCCTACCGGCGCTTCCCAAGTCTCGATTCTAT  619

Qy  728 TTACATTTGAT  738
            |||||||
Db  620 CTTGATTTGGT  630
```

RESULT 15
PCT-US04-23166-352
; Sequence 352, Application PC/TUS0423166
; GENERAL INFORMATION:
; APPLICANT: Purdue Pharma L.P.
; APPLICANT: Jiefei Tong
; APPLICANT: Gang Jin
; APPLICANT: Rui-Ru Ji
; APPLICANT: Yixun Xu
; APPLICANT: Lillian W. Chiang
; APPLICANT: Daniel J. Lavery
; TITLE OF INVENTION: GENES ASSOCIATED WITH RESPONSES TO NEUROPATHIC PAIN
; FILE REFERENCE: 02755/200M584-WOO
; CURRENT APPLICATION NUMBER: PCT/US04/23166
; CURRENT FILING DATE: 2004-07-26
; PRIOR APPLICATION NUMBER: 60/485,101
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 868

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 352
; LENGTH: 2781
; TYPE: DNA
; ORGANISM: Mus musculus
PCT-US04-23166-352

Query Match      11.7%;   Score 136.4;   DB 1;   Length 2781;
Best Local Similarity   56.3%;   Pred. No. 7.2e-21;
Matches  276;   Conservative   0;   Mismatches 211;   Indels    3;   Gaps    1;

Qy    251 AGAGCGAAAGAAAGAAAAAGGAATAGACTAAGTCTTCCTACAGAGGCCTGACTTCCATGGA 310
              ||| || || |  ||||  | |    |  |    ||| |||| | |||  |||  || ||
Db    140 AGTGCAAAGGACATCGAAAACATCGGCTGGGGTTCCTGCTGCGAGAAGTCAGATTCCTGCGAA 199

Qy    311 GAGA--CTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCTCTCCTGAA 367
              |||   |  ||  | | | ||  |  | |||  |  | ||   || | |  | ||
Db    200 CACAGTTCTTCACACAGCAAGAGGACAAAGTAGTAACTTGCCAGAGGGTAAGCCAAGAA 259

Qy    368 GAAGCAGTGAAATGGGCTGAATCATTGACAAATTGCTCTCTCATAGAGATGGAGTGGAT 427
              ||||     ||| ||| || |  | ||||| |  |  | |  ||||  ||| ||
Db    260 GAAGTCAAGAAAATGGGCTGAATCGTTGGAAAAACCTGATTCACCATGAATGTGGACTGGCA 319

Qy    428 GCTTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTGGGTCGCC 487
              ||||| |  | |  |||| |   |     | |||||||   |||| |  |||||||||||
Db    320 GCTTTCAAAGCTTTCCTGAGTCCGAATACAGCGAGGAGAACATTGACTTCTCTGGATCAGC 379

Qy    488 TGTGAAGACTTCAAGAAATGCAAGGAAATCATCATCCTAAAAGCAAAGGCAAATC 547
              ||||| | |  ||| ||| || ||  |  |  | |    | ||  ||| |  |  ||
Db    380 TGTGAGGAGTACAAGAAGATCAAGTCACCTTCTAAACTAAGTCCCAAAGCCAAGAAGATC 439

Qy    548 TATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGTTAACATTGATTTTCATACTAAA 607
              |||||| ||||| | |  | ||  ||  |  || ||| |||||  | ||||  ||| |||
Db    440 TACAACGAGTTCATCTCAGTGCAAGCAACAAAAGAGGTGAACCTGGACTCTTGCACCAGA 499

Qy    608 GAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTTGATACGGCACAAAGC 667
              ||||||   |   | ||  || || || | || |  | ||||  |||| | || || |||
Db    500 GAGGAGACAAGCCGGAACATGTTACAGCCCACAATAACCTGTTTTGATGGAAGCCCAGAAG 559

Qy    668 AGAGTGTACCAGCTCATGGAACATGACAGTTATAAAACGCTTTTGAAATCTGAGACCTAC 727
              ||| |||  ||| |  || |  |  |  | |||||  | | |||| ||| |||| |||
Db    560 AAGATTTTCAACCTGATGGAAAAACGGATTCCTACCGCCGCTTCCTCAGTCTCGATTCTAT 619

Qy    728 TTACATTTGA 737
              ||  |  |||
Db    620 CTTGACCTGA 629
```

Search completed: August 20, 2004, 16:14:19
Job time : 464.364 secs

```
OM nucleic - nucleic search, using sw model

          Copyright  GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

Run on:       August 20, 2004, 05:11:40 ; Search time 3157.02 Seconds
                                          (without alignments)
                                          11010.254 Million cell updates/sec

Title:        US-09-894-749-3
Perfect score: 1164
Sequence:     1 ttttgtaagaaaatctga.........gtactagtggatccgagctc 1164

Scoring table: IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0

Searched:     2751328.9 seqs, 1493109027.6 residues

Total number of hits satisfying chosen parameters:   55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    EST:*
              1:  em_estba:*
              2:  em_esthum:*
              3:  em_estin:*
              4:  em_estmu:*
              5:  em_estov:*
              6:  em_estpl:*
              7:  em_estro:*
              8:  em_htc:*
              9:  gb_est1:*
             10:  gb_est2:*
             11:  gb_htc:*
             12:  gb_est3:*
             13:  gb_est4:*
             14:  gb_est5:*
             15:  em_estfun:*
             16:  em_estom:*
             17:  em_gss_hum:*
             18:  em_gss_inv:*
             19:  em_gss_pln:*
             20:  em_gss_vrt:*
             21:  em_gss_fun:*
             22:  em_gss_mam:*
             23:  em_gss_mus:*
             24:  em_gss_pro:*
             25:  em_gss_rod:*
             26:  em_gss_phg:*
             27:  em_gss_vrl:*
             28:  gb_gss1:*
             29:  gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                         SUMMARIES
```

Result No.	Score	% Query Match	Score	Length	DB	ID	Description
1	1116.4	95.9	1860	11	AK036407	AK036407 Mus muscu
2	605.4	52.1	700	13	BY750625	BY750625 Mus muscu
3	594.4	51.1	1201	9	AL557903	AL557903 Mus muscu
4	592	50.9	655	29	AY420374	AY420374 Mus muscu

No.		Score	Pred				
5	560.8	48.2	666	10	BB660526	BB660526	BB660526
6	557.8	47.9	618	13	BY722727	BY722727	BY722727
7	536.4	46.1	538	13	BU698389	BU698389	LL2in1395
8	512.6	44.0	708	29	AY420372	AY420372	Homo sapi
9	493.8	42.4	616	10	BB629298	BB629298	BB629298
10	476	40.9	661	14	CD466872	CD466872	LeukoN2_8
11	469.6	40.3	723	14	CF181629	CF181629	818478_MA
12	438.6	37.7	764	14	CD468328	CD468328	LeukoS3_2
13	402.8	34.6	810	12	BG564257	BG564257	602586054
14	392.2	33.7	705	9	AV714060	AV714060	AV714060
15	387.8	33.3	626	9	AV714044	AV714044	AV714044
16	387.4	33.3	639	9	AV716055	AV716055	AV716055
17	379.2	32.6	789	14	CD468976	CD468976	LeukoS3_8
18	375.2	32.2	710	10	BB490551	BB490551	BB490551
19	365	31.4	625	9	AV715708	AV715708	AV715708
20	362.8	31.2	589	29	AY420373	AY420373	Pan trogl
21	347.2	29.8	716	14	CD465070	CD465070	LeukoN1_1
22	344.8	29.6	861	14	CF257072	CF257072	pha006_d0
23	335.2	28.8	340	12	BG094619	BG094619	uu92b08.x
24	334.4	28.7	411	13	BY227741	BY227741	BY227741
25	314	27.0	772	13	BU285628	BU285628	60386709
26	308.4	26.5	444	14	CB788440	CB788440	AMGNNUC:N
27	302	25.9	701	9	AV712845	AV712845	AV712845
28	293	25.2	921	13	BX350311	BX350311	BX350311
29	290	24.9	344	13	BY111739	BY111739	BY111739
30	283.4	24.3	344	13	BY213581	BY213581	BY213581
31	280.2	24.1	349	13	BY219601	BY219601	BY219601
32	278.6	23.9	781	13	BU361362	BU361362	603788927
33	275	23.6	351	13	BY111479	BY111479	BY111479
34	270.2	23.2	339	13	BY215437	BY215437	BY215437
35	266	22.9	599	12	BG625427	BG625427	pgn1c.pk0
36	259.4	22.3	536	13	BU469215	BU469215	603363777
37	252.6	21.7	648	28	AZ084304	AZ084304	RPCI-23-3
38	242.8	20.9	610	9	AI981146	AI981146	pat.pk004
39	242.6	20.8	694	13	BU277880	BU277880	603867112
40	240.8	20.7	676	14	CB018381	CB018381	pgn1c.pk0
41	238.4	20.5	632	12	BQ038335	BQ038335	pgn1c.pk0
42	232.2	19.9	641	12	BM491165	BM491165	pgp2n.pk0
43	227.2	19.5	627	12	BQ037232	BQ037232	pgn1c.pk0
44	227	19.5	484	14	N98410	N98410	za69c05.r1
45	215.6	18.5	739	12	BI909559	BI909559	603070761

```
                       ALIGNMENTS

RESULT 1
AK036407
LOCUS        AK036407        1860 bp    mRNA    linear   HTC 19-SEP-2003
DEFINITION   Mus musculus adult male bone cDNA, RIKEN full-length enriched
             library, clone:9830006J10 product:regulator of G-protein signaling
             18, full insert sequence.
ACCESSION    AK036407
VERSION      AK036407.1  GI:26331373
KEYWORDS     HTC; CAP trapper.
SOURCE       Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1
  AUTHORS    Carninci,P. and Hayashizaki,Y.
  TITLE      High-efficiency full-length cDNA cloning
  JOURNAL    Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE    99279253
   PUBMED    10349636
REFERENCE    2
  AUTHORS    Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
             Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE      Normalization and subtraction of cap-trapper-selected cDNAs to
             prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL    Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE    20499374
   PUBMED    11042159
```

REFERENCE   3
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
             Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
             Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
             Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
             Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
             Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
             Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
             sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PUBMED   11076861
REFERENCE   4
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
             FANTOM Consortium.
TITLE   Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
             Group Phase I & II Team.
TITLE   Analysis of the mouse transcriptome based on functional annotation
             of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 1860)
AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
             Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
             Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
             Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
             Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
             Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
             Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
             Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
             Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
             Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
             Takeda,T., Tanaka,T., Tomaru,Y., Toya,T., Yasunishi,A.,
             Muramatsu,M. and Hayashizaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
             RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
             Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
             Please visit our web site for further details.
             URL:http://genome.gsc.riken.go.jp/
             URL:http://fantom.gsc.riken.go.jp/.
FEATURES   Location/Qualifiers
     source   1..1860
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:9830006J10"
                     /db_xref="MGI:2400834"
                     /db_xref="taxon:10090"
                     /clone="9830006J10"
                     /sex="male"
                     /tissue_type="bone"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS   182..889
                     /codon_start=1
                     /note="unnamed protein product; putative
                     regulator of G-protein signaling 18 (MGD|MGI:1927498,
                     GB|NM_022881, evidence: BLASTN, 99%, match=1393)"
                     /protein_id="BAC29417.1"
                     /db_xref="GI:26331374"

                     /translation="MDMSLVFFSQLNMCBSKEKTFFKLMHGSGKEETSIEAKIRAKEK
                     RNRLSLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTR
                     FLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEV
                     IAKSIAQPTLHSFDTAQSRVQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRS
                     FTYNDFQDVKSDVAIWL"
     polyA_signal   1839..1844
                     /note="putative"
     polyA_site   1860
                     /note="putative"
ORIGIN

Query Match   95.9%; Score 1116.4; DB 11; Length 1860;
Best Local Similarity 99.9%; Pred. No. 4.3e-227;
Matches 1117; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Qy   1 TTTTGTAAGAAAATCGAGGAAAGATTCGGGATAGCGGCTTATTCAGGATGTTTTCCT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   49 TTTTTGTAAGAAAATCGAGGAAAGATTCGGGATAGCGGCTTATTCAGGATGTTTTCCT 108

Qy   61 ATGAAATAGCATTCATCTGTGGGAGAGGAGGAGGACTAAGGAAATCTGACATCTGTTGGTC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   109 ATGAAATAGCATTCATCTGTGGGAGAGGAGGAGGACTAAGGAAATCTGACATCTGTTGGTC 168

Qy   121 ACTGGGACAGAATATGATAGTCACTGGTTTCTTCTTCAATTAAATATGTGTGAATC 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   169 ACTGGGACAGAATATGATAGTCACTGGTTTCTTCTTCTCAATTAAATATGTGTGAATC 228

Qy   181 AAAAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGGAAAGAGAAGAAACAAGCATCGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   229 AAAAGAGAAAACTTTTTCAAACTAATGCATGGGTCAGGGAAAGAGAAGAAACAAGCATCGA 288

Qy   241 GGCCAAAATCAGAGCGAAAGAAAAGAATAGACTAAGTCTTCTCCTACAGAGGCCTGA 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   289 GGCCAAAATCAGAGCGAAAGAAAAGAATAGACTAAGTCTTCTCCTACAGAGGCCTGA 348

Qy   301 CTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCCTTGGCCAAGAAACAAGAGTCTC 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   349 CTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCCTTGGCCAAGAAACAAGAGTCTC 408

Qy   361 TCCTGAGAGCAGTGAAATGGCTGAATCATTGACAAATTGCTCTCTCATAGAGATGG 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   409 TCCTGAGAGCAGTGAAATGGCTGAATCATTGACAAATTGCTCTCTCATAGAGATGG 468

Qy   421 AGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGACATTGAATTTTG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   469 AGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGACATTGAATTTTG 528

Qy   481 GGTCGCCTGTGAAGACTTCAAGAAATGCAAGAACCTCAACAATCATCCTAAAAGCAAA 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   529 GGTCGCCTGTGAAGACTTCAAGAAATGCAAGAACCTCAACAATCATCCTAAAAGCAAA 588

Qy   541 GGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAGAGGTTAACATTGATTTTCA 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   589 GGCAATCTATGAGAAATTCATTCAGAATGATGCCCCAAGAGGTTAACATTGATTTTCA 648

Qy   601 TACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTGATACGGC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   649 TACTAAAGAAGTAATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTGATACGGC 708

Qy   661 ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAAATCTGA 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   709 ACAAAGCAGAGTGTACCAGCTCATGGAACATGACAGTTATAAACGCTTTTTGAAATCTGA 768

Qy   721 GACCTACTTACATTTGATAGAGGAAGACCTCAGAGACAAACCTTAGGAGACGATC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   769 GACCTACTTACATTTGATAGAGGAAGACCTCAGAGACAAACCTTAGGAGACGATC 828

Qy   781 ACGATCATTTACTTACAATGATTTCCAAGATGTAAAGTCAGATGTTGCCATTGGTTATG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   829 ACGATCATTTACTTACAATGATTTCCAAGATGTAAAGTCAGATGTTGCCATTGGTTATG 888

Qy   841 AGTAAAAGTCATTTGTCTTCTTTGATAGTGTATATCTAAAATATATACTAATA 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   889 AGTAAAGTCATTGTCTTCTTTTGATAGTGTGATGTGTATATCTAAAATATATACTAATA 948
```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Qy   901 CTAATGTGTACTTCTAAAATATAGCTTGTGTGTATAAGAGAGATGATTCATTTTTAAAAT 960
```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db   949 CTAATGTGTACTTCTAAAATATAGCTTGTGTGTATAAGAGAGATGATTCATTTTTAAAAT 1008

Qy   961 ACACCATGCAAATACATATTAAATGTAAGAACTTTTATATTATATATATAATTCATC 1020
```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db   1009 ACACCATGCAAATACATATTAAATGTAAGAACTTTTTATATTATATATACTAAAATAATTCATC 1068

Qy   1021 ATCTATCTTCCGAAATATTTATGAAAATCTATCTGATATTCTATTCTAATAAAATTCTT 1080
```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db   1069 ATCTATCTTCCGAAATATTTTATGAAAATCTATCTGATATTCTATTCTAATAAAATTCTT 1128

Qy   1081 TATTTCTACAATACAGTCAGTAAGAAGAGCTTTGAA 1118
```
                |||||||||||||||||||||||||||||||||||
```
Db   1129 TATTTCTACAATAACAGTCAGTAAGAAGAAGCTTTGTA 1166

RESULT 2
BY750625
LOCUS   BY750625   700 bp   mRNA   linear   EST 17-DEC-2002
DEFINITION   BY750625 RIKEN full-length enriched, activated spleen Mus musculus
             cDNA clone F830007B06 5', mRNA sequence.
ACCESSION   BY750625
VERSION   BY750625.1   GI:27180611
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 700)
AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
             Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
             Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
             Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
             Hume,D.A., Quackenbush,J., Schrim,L.M., Kanapin,A., Matsuda,H.,
             Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
             Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
             Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
             Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
             Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
             Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
             Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
             Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
             Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
             Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
             Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
             Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
             Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
             Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
             Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
             Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
             Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
             Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
             Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
             Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
             Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
             Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE   Analysis of the mouse transcriptome based on functional annotation
             of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
MEDLINE   22354683
PUBMED   12466851
COMMENT   Contact: Yoshihide Hayashizaki
             Laboratory for Genome Exploration Research Group, RIKEN Genomic
             Sciences Center(GSC) Yokohama Institute
             The Institute of Physical and Chemical Research (RIKEN)
             1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
             Tel: 81-45-503-9222
             Fax: 81-45-503-9216
             Email: genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in RIKEN.
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES   Location/Qualifiers
     source   1..700
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="NOD"
                     /db_xref="taxon:10090"
                     /clone="F830007B06"
                     /tissue_type="activated spleen"
                     /clone_lib="RIKEN full-length enriched, activated spleen"

ORIGIN

Query Match   52.0%;   Score 605.4;   DB 13;   Length 700;
Best Local Similarity   97.0%;   Pred. No. 1.9e-118;
Matches   623;   Conservative   0;   Mismatches 18;   Indels 18;   Gaps 1;

Qy   1 TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTCAGGATGTTTTCCT 60
```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db   60 TTTTTGTAAGAAAAATCTGAGGAAAGATTCGGGATAGCGCTTTATTCAGGATGTTTTCCT 119

Qy   61 ATGAAATAGCATTCATCTGTGGGAGAGAAGAAGACTAAGGAAATCTGACATCTGTTGGTC 120
```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db   120 ATGAAATAGCATTCATCTGTGGGAGAGAAGAAGACTAAGGAAATCTGACATCTGTTGGTC 179

Qy   121 ACTGGGACAGAATATGGATATGTCACTGGTTTTCTTCCTCTCAATTAAATATGTGTGAATC 180
```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db   180 ACTGGGACAGAATATGGATATGTCACTGGTTTTCTTCCTCTCAATTAAATATGTGTGAATC 239

Qy   181 AAAAGAGAAAACTTTTTCAAACTAATGCATGGTCAGGGAAGAAGAAGAAGAAGCATCGA 240
```
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db   240 AAAAGAGAAAACTTTTTCAAACTAACTAATGCATGGTCAGGGAAGAAGAAGAAGAAGCATCGA 299

Qy   241 GGCCAAATCAGAGCGAAGAAGCCGAAAGAAGAAATAGCAATAGCTAAGTCTTCCTCCACGAGGCCTGA 300
```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db   300 GGCCAAATCAGAGCGAAGAAGCCGAAAGAAGAAATAGACTAAGTCTTCCTCCACGAGGCCTGA 359

Qy   301 CTTCCATGGAGACTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCTC 360
```
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db   360 CTTCCATGGAGACTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCTC 419

Qy   361 TCCTGAAGAAGCAGTGAAATGGCTGAATCATTTGACAAATTGCTCTCCTCATAGAGATGG 420
```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db   420 TCCTGAAGAAGCAGTGAAATGGCTGAATCATTTGACAAATTGCTCTCCTCATAGAGATGG 479

Qy   421 AGTGGATGCTTTACCAGATTTCCTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG 480
           |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   480 AGTGGATGCTTTTACCAGATTTCCTAAGACTGAATTCAGTGAGGAGAACATTGAATTTTG 539

Qy   481 GGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCCAACAAATCATCCTAAAAGCAAA 540
           ||||||||||| |||||||||||| ||||||||||||||||| |||||||||||||||||
Db   540 GGTCGCCTGTGAAGACTTCNAGANATGCAGGAACCTCCAACAATCATCCTAAAAGCAAG 599

Qy   541 GGCAATCTATGAGAAATTCATTCAGAAATGATGCCCCAAAGAGGTTAACATTGATTTTCA 600
           ||| ||||||||||||||||||||| |||||||||| ||||||||||||||||||||||
Db   600 GGC-ATCTATGAGAAATTCATTCAGAAATGATGCCCNCAAAGAGGTTAACATTGATTTTCA 658

Qy   601 TACTAAAGAAGTAATTGCTAAGAGCATGCCCCAGCCCACTCT 642
           ||| ||||||||||| ||||||||| |||||| ||||||||||
Db   659 TACTANAGAAGTATTNNGCTAGAGCATCGNCCCCAGCCCACTCT 700

RESULT 3
AL557903
LOCUS   AL557903   1201 bp   mRNA   linear   EST 31-MAY-2003
DEFINITION AL557903 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
            Homo sapiens cDNA clone CS0DJ003YJ10 5-PRIME, mRNA sequence.
ACCESSION   AL557903
VERSION   AL557903.2 GI:31279703
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201).
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT   On Feb 15, 2001 this sequence version replaced gi:12901962.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 261.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DJ003DE05QP1&cluster=261.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fullength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DJ003DE05QP1.
FEATURES   Location/Qualifiers
     source   1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ003YJ10"
                     /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                     /cell_line="JURKAT"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
                     10-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match   51.1%; Score 594.4; DB 9; Length 1201;
Best Local Similarity   81.0%; Pred. No. 4.1e-116;
Matches 714; Conservative   2; Mismatches 163; Indels   3; Gaps   2;

Qy   4 TTGTAAGAAAAATCGAGGAAGAATTCGGGATAGCGCCTTATTC-AGGATGTTTTCCTAT   62
           |||| |||| ||||||||||||||| ||| |||||||| ||||| |||||||||||||||
Db   80 TTTTTGTAACATTACTGTAAGAGTTGTGATAACTTTTTATTCTACTACTATATGTAT 139

Qy   63 GAAATAGCATTCATCTGTGGGAGAGAAGGACTAAGAAAATCGACATCTGTTGGTCAC 122
           ||||||||||||||| ||||||||||||||||||||| |||||||||||||||||||
Db   140 GGAAWAGTATTAAWAAATGAACTAGGGAAGGATGTAATAAAATTAGACATCTCTTCATT-- 197

Qy   123 TGGGACGAAATATGGATATGTCACTCGTTTCTCTCCAATTAAATATGTGAATCAA 182
           || |||||||||| ||||| |||| |||||||||||||||||||| |||||||||
Db   198 TTAGAGAAGATGAAACACATTGCTTTTTCTTTTTCTCAAATAAATATGTGAATCAA 257

Qy   183 AAGAGAAAACTTTTTCAAGCTCATGGGTCAGGGAAAGAAGAAGAACAAGCATCGAGG 242
           |||||||||||||||||||| ||||||||| ||||||||||| ||||||||||||||
Db   258 AAGAAAAACTTTTTCAAGTTAATACATGGTTCAGGAAAAGAAGAAGAACAAGCAAGAAG 317

Qy   243 CCAAAATCAGAGCGAAAGAAATAGAACTAAGTCTTCTCCTACAGAGGCCTGACT 302
           ||||||||||||| ||||||||| |||||||| |||||||||||||||||||||
Db   318 CCAAAATCAGAGCTAAGGAAAAAAGAGAAAATAGAACTAAGTCTTCTTCTGTGCAGAAACCTGAGT 377

Qy   303 TCCATGAGAGACTCAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCTCTC 362
           |||||||||||||||||| ||||||||| ||| |||||||||||||||||||| ||||
Db   378 TTCATGAAGACACCCGCYCCAGTAGATCTCGGGCACTTGGCCAAAGAAGAACAAGAGTCTCCC 437

Qy   363 CTGAAGAAGCAGTGAATGGGCTGAATCATTGACAAATGCTCTCTCCATAGAGATGGAG 422
           |||||| |||||||| |||||||||||||| ||||||||||||| ||||||||||||||
Db   438 CTGAAGAGCAGTGAAATGGGGTGAATCATTGACAAACTGCTTTCCCATAGAGATGGAC 497

Qy   423 TGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTGGG 482
           |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   498 TAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAAGAAATATTGAATTTTGGA 557

Qy   483 TCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAAGG 542
           ||||||||||||||||||||||||||||| |||||||||||| |||||||||||||| |||
Db   558 TAGCCTGTGAAGATTTCAAGAAAAGCAAGGAACCTCAACAAATTCACCTTAAAAGCAAAAG 617

Qy   543 CAATCTATGAGAAATTCATTCAGACGATGATGCCCCAAAGAGAGGTTAACATTGATTTTCATA 602
           ||||| |||||||||||||| || |||||||||| ||||||||||||||||||||||| ||
Db   618 CAATATATGAGAAATTTATACAGACTGATGCCCCAAAAAGAGAGGTTAACCTTGATTTTCACA 677

Qy   603 CTAAAGAAGTAATTGCTAAGAGCATCGCCCCACTCCTCCCACAGTTTTGATACGGCAC 662
           ||||||| ||||| |||||||| ||||||||| |||||||||||||||||||| ||||||
Db   678 CAAAAGAAGTCATTACAAACAGCATCCATCATCACTCCACTACCCTCCACAGTTTTGATGCTGCAC 737

Qy   663 AAAGCAGAGTGTACCAGCTCATGGGAACCATGACCATGACAGTTATAAACGCTTTTGAAACTGAGA 722
           |||||||||||| |||||||||| |||| |||||||||||||||||||| |||| ||||||||||
Db   738 AAAGCAGAGTGATCGTCTCATGGGACCAACAGCAGTTATACACGTTTTCTGAAATCTGACA 797

Qy   723 CCTACTTACATTTGATAGAAGGAAGACCTTCAGAGACCAACAAACCTTAGGAGAGAGACGATCAC 782
           ||| |||||||||||||||| ||||||||||||||||| |||||||||| ||||||||||||
Db   798 TCTATTTAGACTTGATGGAAGGAAGACCTCAGAGAACCAACCAACAAATCTTAGGAGAGAGACCATCAC 857

Qy   783 GATCATTTACTTACAATGATTTCCAAGATGTAAAGTCAATCAGATGTTGGCCATTTGGTTATGAG 842
           |||||||||||||||||||||||||||||||||||||||| ||||||| |||||||||||||||
Db   858 GCTCATTTTACCTGCAATGAATTCCAAGATGTCAATCAGATGTTGGCCATTTGGTTATAAA 917

Qy   843 TAAAAGTCATTTGTCTTCTTTTGATAGTGATATGTGTATATCT 884
           |||||||||||||| ||||||| |||||||||| ||||||||||
Db   918 GAAAATTGATTTTGCTCATTTTTATGACAAACTTATATACATCT 959

RESULT 4
AY420374
LOCUS   AY420374   655 bp   DNA   linear   GSS 12-DEC-2003
DEFINITION Mus musculus HCM7202 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY420374
VERSION   AY420374.1 GI:39776331
KEYWORDS   GSS.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 655)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios

RESULT 5

LOCUS   BB660526   666 bp   mRNA   linear   EST 26-OCT-2001
DEFINITION BB660526 RIKEN full-length enriched, 13 days embryo stomach Mus
            musculus cDNA clone D530016E22 5', mRNA sequence.
ACCESSION   BB660526
VERSION   BB660526.1 GI:16494345
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 666)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
  TITLE   RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/,
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
            wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,Y., Kira,A.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
            RIKEN integrated sequence analysis (RISA) system-384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
            Sugahara,Y. and Hayashizaki,Y.
            Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
            Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
            Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
            Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
            Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
            e mouse tissues.
FEATURES   Location/Qualifiers
     source   1..666
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="D530016E22"
                     /tissue_type="stomach"
                     /dev_stage="13 days embryo"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, 13 days embryo
                     stomach"
                     /note="Site 1: SalI; Site 2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken

JOURNAL   Science 302 (5552), 1960-1963 (2003)
PUBMED   14671302
REFERENCE   2 (bases 1 to 655)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering them
            based on alignment.
FEATURES   Location/Qualifiers
     source   1..655
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene   <1..>655
                     /locus_tag="HCM7202"
ORIGIN

Query Match   50.9%; Score 592; DB 29; Length 655;
Best Local Similarity   92.5%; Pred. No. 1.4e-115;
Matches 655; Conservative   0; Mismatches   0; Indels   53; Gaps   1;

Qy   134 ATGGATATGTCACTGGTTTCTTCTCTCAATTAAATATGTGAATCAAAAGAGAAAACT 193
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGGATATGTCACTGGTTTCTTCTCTCAATTAAATATGTGAATCAAAAGAGAAAACT 60

Qy   194 TTTTTCAAACTAATGCATGGGTCAGGGAAGAAGAAGAAGAAA 253
          ||||||||||||||||||||||||||||||||||||||||||
Db   61 TTTTTCAAACTAATGCATGGGTCAGGGAAGAAGAAGAAGAAA 120

Qy   254 GCGAAAGAAAAAAGGAATAGACTAAGCTCTTCTCCTACAGAGGGCCTGACTTCCATGGAGAG 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GCGAAAGAAAAAAGGAATAGACTAAGCTCTTCTCCTACAGAGGGCCTGACTTCCATGGAGAG 180

Qy   314 ACTCAAGCCAGTAGATCTGCCCCTCTTGGCCAAAGAAACAAGAGTCTCTCCTGAAGAAGCA 373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACTCAAGCCAGTAGATCTGCCCCTCTTGGCCAAAGAAACAAGAGTCTCTCCTGAAGAAGCA 240

Qy   374 GTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGATGGGAGTGGATGCTTTT 433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCATAGAGATGGGAGTGGATGCTTTT 300

Qy   434 ACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGGGTCGCCTGTGAA 493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGGGTCGCCTGTGAA 360

Qy   494 GACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAAGGCAATCTATGAG 553
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GACTTCAAGAAATGCAAGGAACCTCAACAAATCATCC----------------------- 397

Qy   554 AAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTCATACTAAAGAAGTA 613
                                                ||||||||||||||||||||||||
Db   398 -------------------------------------GTTAACATTGATTTCATACTAAAGAAGTA 427

Qy   614 ATTGCTAAGAGCATCGCCCAGCCACTCTTCCACAGTTTTGATACGGCCAAAAGCAGAGTG 673
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 ATTGCTAAGAGCATCGCCCAGCCACTCTTCCACAGTTTTGATACGGCCAAAAGCAGAGTG 487

Qy   674 TACCAGCTCATGGAACATGACAGTTATAAACGCTTTTGAAAATCGAGACCTACTTACAT 733
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 TACCAGCTCATGGAACATGACAGTTATAAACGCTTTTGAAAATCGAGACCTACTTACAT 547

Qy   734 TTGATAGAAGGAAGACCTCAGAGACCAACAAACCTTAGGAGACGATCACGATCATTTACT 793
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 TTGATAGAAGGAAGACCTCAGAGACCAACAAACCTTAGGAGACGATCACGATCATTTACT 607

Qy   794 TACAATGATTTCCAAGATGTAAAGTCAGATGTTGCCATTGGTTATGA 841
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   608 TACAATGATTTCCAAGATGTAAAGTCAGATGTTGCCATTGGTTATGA 655

contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGGATTCGAGTTATTAATTAATTAATCCCCCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

REFERENCE	1 (bases 1 to 618)
AUTHORS	Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
Chothia,C., Corbani,L.E., Cousins,S.A., Dalla,E., Dragani,T.A.,
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKenzie,L., McKenzie,L., Miki,H., Nagashima,T.,
Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE	Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	22354683
PUBMED	12466851
COMMENT	Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
FEATURES	Location/Qualifiers
source	1..618
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"

Query Match   48.2%;   Score 560.8; DB 10; Length 666;
Best Local Similarity   95.6%; Pred. No. 5.8e-109;
Matches 585; Conservative   0; Mismatches   26; Indels   1; Gaps   1;

```
Qy     1 TTTTGTAAGAAAATCTGAGGAAAGATTCGGGATAGCCGCCTTTATTCAGGATGTTTTCCT 60
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    55 TTTTGTAAGAAAATCTGAGGAAAGATTCGGGATAGCCGCCTTTATTCAGGATGTTTTCCT 114

Qy    61 ATGAAATAGCATTCATCTGTGGAGAGAGAGGACTAAGGAAATCTGACATCTGTTGGTC 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   115 ATGAAATAGCATTCATCTGTGGAGAGAGAGGACTAAGGAAATCTGACATCTGTTGGTC 174

Qy   121 ACTGGGACAGAATATGGATATGTCACTGGTTTCTTCTCCTCAATTAAATATGTGTGAATC 180
           ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   175 ACTGGGACAGAATATGGATATGTCACTGGTTTCTTCTCCTCAATTAAATATGTGTGAATC 234

Qy   181 AAAAGAGAAAAC-TTTTTTCAAACTAATGCATGGTCAGGGAAGAAGAGAAACAAGCATCG 239
           ||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||
Db   235 AAAAGAGAAAACTTTTTTTCAAACTAATCATGGGTCAGGGAAGAAGAGAAACAAGCATCG 294

Qy   240 AGGCCAAAATCAGAGCGAAAGAAAGAAAAAAGGAATAGACTAAGTCTTCTCCTACAGAGGCCTG 299
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   295 AGGCCAAAATCAGAGCGAAAGAAAGAAAAAAGGAATAGACTAAGTCTTCTCCTACAGAGGCCTG 354

Qy   300 ACTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCTCTGGCCAAAGAAACAAGAGTCT 359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   355 ACTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCT 414

Qy   360 CTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTTCTTCATAGAGATG 419
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   415 CTCCTGAAGAAGCAGTGAAATGGGCTGAATCATTTGACAAATTGCTTCTTCATAGAGATG 474

Qy   420 GAGTGGATGCTTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTT 479
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   475 GAGTGGATGCTTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTT 534

Qy   480 GGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAA 539
           |||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db   535 GGGTCGCCTGTGGAGACTTCAAGAAATGCCAGGAACCTCAACAAATCATCCCCCTAAAGCA 594

Qy   540 AGGCAATCTATGAGAAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTTC 599
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   595 AGGGCATCTATGAGAAAATTCATTCAAAAAAGGGAGCCCCAAAGAGGGTAACANNTGAINTC 654

Qy   600 ATACTAAAGAAG 611
           ||||| |||||
Db   655 ATACTANAAAAG 666
```

RESULT 6
BY722727
LOCUS   BY722727   618 bp   mRNA   linear   EST 17-DEC-2002
DEFINITION BY722727 RIKEN full-length enriched, adult male bone Mus musculus
            cDNA clone 9830112K19 5', mRNA sequence.
ACCESSION   BY722727
VERSION   BY722727.1 GI:27135844
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

RESULT 7
BU698389
LOCUS   BU698389   538 bp   mRNA   linear   EST 09-OCT-2002

DEFINITION   LL2in1395lT7 Hematopoietic Stem Cell Subtracted Library Mus
             musculus cDNA 5' similar to RGP5, regulator of G-protein
             signalling, mRNA sequence.
ACCESSION   BU698389
VERSION   BU698389.1 GI:23611104
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 538)
AUTHORS   Phillips,R.L., Ernst,R.E., Brunk,B.P., Ivanova,N., Mahan,M.A.,
             Deanehan,J.K., Moore,K.A., Overton,G.C. and Lemischka,I.R.
TITLE   The genetic program of hematopoietic stem cells
JOURNAL   Science 288 (5471), 1635-1640 (2000)
MEDLINE   20295303
PUBMED   10834841
COMMENT   Contact: Lemischka, Ihor R.
             Department of Molecular Biology
             Princeton University
             Lewis Thomas Laboratory, Washington Road, Princeton, NJ 08544, USA
             Tel: 609 258 2838
             Fax: 609 258 2759
             Email: ilemischka@molbio.princeton.edu
             These ESTs are derived from a subtracted cDNA library enriched for
             gene products expressed in day 14-14.5 fetal liver hematopoietic
             stem cells defined as Lineageneg/lo, AA4.1pos, cKitpos, Ly6A/E
             (Sca-1)pos
             seq primer: M13Reverse or T7.
FEATURES   Location/Qualifiers
     source   1..538
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /tissue_type="Hematopoietic"
                     /cell_type="Stem Cells; Lineageneg/lo, AA4.1pos, cKitpos,
                     Ly6A/E (Sca-1)pos"
                     /dev_stage="Embryonic day 14-14.5"
                     /lab_host="DH10B"
                     /clone_lib="Hematopoietic Stem Cell Subtracted Library"
                     /note="Organ: Fetal Liver; Vector: Sport 1; Site_1: Sal I;
                     Site_2: Not I; Two directionally cloned cDNA libraries
                     were made from fetal liver hematopoietic stem cells
                     enriched to be Lineageneg/lo, AA4.1pos, cKitpos, Ly6A/E
                     (Sca-1)pos and from AA4.1neg fetal liver
                     cells. Subtractive hybridization was performed by
                     hybridization of the target, Scapos, single stranded cDNA
                     library in pSport1 to biotinylated RNA transcribed from
                     the driver, AA4.1neg cDNA library in pSport2 with inserts
                     cloned in the complementary orientation. For detailed
                     protocols and additional information please see our
                     website at http://stemcell.princeton.edu."

ORIGIN

Query Match   46.1%; Score 536.4; DB 13; Length 538;
Best Local Similarity   99.8%; Pred. No. 9.1e-104;
Matches 537; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   275 CTAAGTCTTCTCCTACAGAGGCCTGACTTCCATGGAGGAGACTCAAGCCAGTAGATCTGCC 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CTAAGTCTTCTCCTACAGAGGCCTGACTTCCATGGAGGAGACTCAAGCCAGTAGATCTGCC

Qy   335 CTCTTGGCCAAAGAAACAAGAGTCTCCTGAAGAAGCAGCAGTGAAATGGGCTGAATCATT 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTCTTGGCCAAAGAAACAAGAGTCTCCTGAAGAAGCAGCAGTGAAATGGGCTGAATCATT 120

Qy   395 GACAAATTGCTCTCTCATAGAGATGGAGTGGATGCTTTTACCAGATTTCTTAAAACTGAA 454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GACAAATTGCTCTCTCATAGAGATGGAGTGGATGCTTTTACCAGATTTCTTAAAACTGAA 180

Qy   455 TTCAGTGAGGAACACATTGAATTTGGTCGCCTGTGAAGACTTCAAGAAATGCAAGGAA 514
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

ORIGIN

Query Match   47.9%; Score 557.8; DB 13; Length 618;
Best Local Similarity   99.3%; Pred. No. 2.5e-108;
Matches 559; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   1 TTTTTGTAAGAAAATCTGAGGAAGATCTCGGATAGCCTTATTCAGGATGTTTCCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   56 TTTTTGTAAGAAAATCTGAGGAAGATCTCGGATAGCCTTATTCAGGATGTTTCCT 115

Qy   61 ATGAAATAGCATTCATCTGTGGGAGAGAAGGACTAAGGAAATCGACATCTGTTGGTC 120
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116 ATGATATAGCATTCATCTGTGGGAGAGAAGGACTAAGGAAATCGACATCTGTTGGTC 175

Qy   121 ACTGGGACAGAATATGGAATATGTCACTGGTTTCTTCTCTCAATTAAATATGTGAATC 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   176 ACTGGGACAGAATATGGAATATGTCACTGGTTTCTTCTCTCAATTAAATATGTGAATC 235

Qy   181 AAAAGAGAAAACTTTTTCAAACTAATGCATGGTCAGGGAAAGAAGAAACAAGCATCGA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   236 AAAAGAGAAAACTTTTTCAAACTAATGCATGGTCAGGGAAAGAAGAAACAAGCATCGA 295

Qy   241 GGCCAAAATCAGAGCGCGAAAGAAAAAGGAATAGACTAAGTCTTCTTCCTACAGAGGCCTGA 300
              ||||||||||||||||| ||||||||||||||||||||||||||||| |||||||||||||
Db   296 GGCCAAAATCAGAGCGGGAAAGAAAAAGGAATAGACTAAGTCTTCT TCCTACAGAGGCCTGA 355

Qy   301 CTTCCATGGAGGAGAGACTCAAGCCTAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCTC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   356 CTTCCATGGAGGAGAGACTCAAGCCTAGTAGATCTGCCCTCTTGGCCAAAGAAACAAGAGTCTC 415

Qy   361 TCCTGAAGAAGCAGTGAAATGGGCTGAAATCATTTGACAAATTGCTCTCTCATAGAGATGG 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   416 TCCTGAAGAAGCAGTGAAATGGGCTGAAATCATTTGACAAATTGCTCTCTCATAGAGATGG 475

Qy   421 AGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGACATTGAATTTTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   476 AGTGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGGAGACATTGAATTTTG 535

Qy   481 GGTCGCCTGAGACTTCAAGAAGTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAA 540
              |||||||||||||||||||||||||||||||||||||||||||| |||| |||||||||||| |||
Db   536 GGTCGCCTGAGACTTCAAGAAGTTCAAGAAATGCAAGGAACCTCNACANATCATCCTAAAAGGCAAA 595

Qy   541 GGCAATCTATGAGAAATTCATTC 563
              |||||||||||||||||||||||
Db   563 GGCAATCTATGAGAAATTCATTC

Qy   596 GGCAATCTATGAGAAATTCATTC 618
              |||||||||||||||||||||||
Db   618

```
Db   181 TTCAGTGAGGAGAACATTGAATTTTGGGTCGCCTGTGAAGACTTCAAGAAATCAAGGAA 240

Qy   515 CCTCAACAAATCATCCTAAAAGCAAAGCAATCTATGAGAAATTCATTCAGAATGATGCC 574
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCAACAAATCATCCTAAAAGCAAAGGCAATCATGAGAAATTCATTCAGAATGATGCC 300

Qy   575 CCCAAAGAGGTTAACATTGATTTCATACTAAAGAAGTAATTGCTAAGAGCATCGCCCAG 634
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCCAAAGAGGTTAACATTGATTTTCATACTAAAGAAGTAATTGCTAAGAGCATCGCCCAG 360

Qy   635 CCCACTCTCCACAGTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGAACATGAC 694
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CCCACTCTCCACAGTTTGATACGGCACAAAGCAGAGTGTACCAGCTCATGGAACATGAC 420

Qy   695 AGTTATAAACGCTTTTGAAATCTGAGACCTACTACATTTGATAGAGGAAGACCTTCAG 754
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGTTATAAACGCTTTTGAAATCTGAGACCTACTACATTTGATAGAGGAAGACCTTCAG 480

Qy   755 AGACCAACAAACCTTAGGAGACGATCACGATCA TTCACTACACAATGATTTCCAAGATG 812
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AGACCAACAAACCTTAGGAGACGATCACGATCA TTCACTACACAATGATTTCCAAGATG 538

RESULT 8
AY420372
LOCUS       AY420372              708 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Homo sapiens HCM7202 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY420372
VERSION     AY420372.1  GI:39776329
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 708)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 708)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering them
            based on alignment.
FEATURES             Location/Qualifiers
     source          1..708
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1..>708
                     /locus_tag="HCM7202"
ORIGIN

Query Match             44.0%;  Score 512.6;  DB 29;  Length 708;
Best Local Similarity   79.9%;  Pred. No. 1e-98;
Matches  566;  Conservative    0;  Mismatches 142;  Indels    0;  Gaps    0;

Qy   134 ATGGATATGTGCACTGGTTTTCTTCTTCTCAATTAAATATGTGAATCAAAAGAGAAAACT 193
             |  ||||| |||||  |||| ||||| ||||   || ||||| |||||||||||||||||
Db     1 ATGGAAACAACAATTGCTTTTCTTTTCTTCAAATAAATAATATGTGTGAATCAAAAGAAAAAACT 60

Qy   194 TTTTTCAAACTAATGCATGGGTCAGGGAAAGAAGAAGAAGAAGCATCGAGCGCCAAAATCAGA 253
             ||||||    ||  ||||||| |||||||||
```

```
Db    61 TTTTTCAAGTAATACATGATGGTTCAGGAAAGAAGAAGAAGCAAAATCAGA 120

Qy   254 GCGAAAGAAAAAGGAATAGACTAAGCTTCTTCCCACAGAGCCTGACTTCCATGGAGAG 313
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GCTAAGGAAAAAGAAATAGACTAAGCTTCTTGTGCAGAAACCTGAGTTTCATGGAGAC 180

Qy   314 ACTCAAGCCAGTAGATCTGCCCTCTTGGCCAAAGAACAAGAGTCTCTCCGAAGAAGCA 373
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCCGCTTCCAGTAGATCTGGGCACTTGGCCAAAGAACAAGAGTCTCCCCTGAAGAGGCA 240

Qy   374 GTGAAATGGGCTGAATCATTTGACAAATTGCTCTCTCCATAGAGATGGAGTGGATGCTTTT 433
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTGAAATGGGGTGAATCATTTGACAAACTGCTTCCCATAGAGATGGACTAGAGGCTTTT 300

Qy   434 ACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAACAATGAATTTGGGTCGCCTGTGAA 493
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ACCAGATTTCTTAAAACTGAATTCAGTGAGGAGAAAATATTGAATTTTGGATGCCTTGAA 360

Qy   494 GACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAAGGCAATCTATGAG 553
             |||||||||||||||||||||||||||||||||
Db   361 GATTTCAAGAAAGAAAGCAAGGGACCTCAACAAATTCACCNNNNNNNNNNNNNNNNNN 420

Qy   554 AAATTCATTCAGAATGATGCCCCCAAAGAGGTTAACATTGATTTTCACACAAAAGAAGTA 613
Db   421 NNNNNNNNNNNNNNNNNNNNNNNNNGTAACCTTGATTTTCACA 480

Qy   614 ATTGCTAAGAGCATCGCCCAGCCCACTCTCCACAGTTTGATACGGCACAAAGCAGAGTG 673
             |||||||||||||||||||||||||||||||
Db   481 ATTACAACACCATCACTCAACCTCCACACCCTCCACAGTTTGATGCGTGCACAAAGCAGAGTG 540

Qy   674 TACCAGCTCATGGAACATGACAGTTATAAACGCTTTTGAAATCTGAGAATCTACTTACAT 733
             |||||||||||||||||||||||||||||||||
Db   541 TATCAGCTCATGGAACAAGACAGTTATACACGTTTTCTGAAATCTGACATCTATTAGAC 600

Qy   734 TTGATGAAGGAAGACCTCAGAGACCAACAAACCTTAGGAGACGATCACGATCATTTACT 793
             |||||||||||||||||||||||||||||||
Db   601 TTGATGGAAGGAAGACCTCAGAGACCAACAAATCTTAGGAGACGATCACGCTCATTTACC 660

Qy   794 TACAATGATTCCAAGATGTAAAGTCAGATGTGCCATTGGTTATGA 841
             |||||||||||||||||||||||||||||||
Db   661 TGCAATGAATTCCAAGATGTACAATCAGATGTGCCATTTGGTTATAA 708

RESULT 9
BB629298
LOCUS       BB629298                616 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB629298 RIKEN full-length enriched, adult male bone Mus musculus
            cDNA clone 9830006I10 5', mRNA sequence.
ACCESSION   BB629298
VERSION     BB629298.1  GI:16466502
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 616)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp.
```

```
Db  226  AAAGAGAAAACTTTTTCAAACTAATGCATGGTCATGGTCAGGGAAGAAGAACAGCATCGA  285
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241  GGCCAAATCAGAGCGAAAGAAAAAAGGAATAGACTAAGTCTTCCTCCACAGAGGCCTGA  300
         ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db  286  GGCCAAATCAGAGCGAAAGAAAAAAGGAATAGACTAAGTCTTCCTCCTACAGAGGCCTGA  345
Qy  301  CTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCTCTCTGCCCAAAGAAACAAGAGTCTC  360
         |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  346  CTTCCATGGAGAGACTCAAGCCAGTAGATCTGCCCTCTCTGGCCAAAGAAACAAGAGTCTC  405
Qy  361  TCCTGAAGAAGCAGTGAAATGGGCTGAATCATTGACAAAATTGCTCTCTCATAGAGATGG  420
         ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  406  TCCTGAAGAAGCAGTGAAATGGGCTGAATCATTGACAAAATTGCTTTCTCATAGAGATGG  465
Qy  421  AGTGGATGCTGTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  466  AGTGGATGCTGTTTTACCAGATTCTTAAAACTGAATTCAGTGAGGAGAACATTGAATTTTG  525
Qy  481  GGTCGCCTGTGAAGACTTCAAGAAA  505
         ||| ||| ||||||||||| |||||
Db  526  GGGTCCCTGTGAAGACTTTAAAGAA  550
```

RESULT 10

```
CD466872
LOCUS       CD466872                       661 bp    mRNA      linear   EST 04-JUN-2003
DEFINITION  LeukoN2_8_B10.g1_A024 Unstimulated peripheral blood leukocytes N2
            Equus caballus cDNA clone LeukoN2_8_B10_A024 5', mRNA sequence.
ACCESSION   CD466872
VERSION     CD466872.1  GI:31388140
KEYWORDS    EST.
SOURCE      Equus caballus (horse)
  ORGANISM  Equus caballus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE   1  (bases 1 to 661)
  AUTHORS   Vandenplas,M., Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S.,
            Moore,J.N., Liang,C., Sun,F., Sullivan,R., Shah,M. and Pratt,L.H.
  TITLE     An EST database from equine (Equus caballus) unstimulated
            peripheral blood leukocytes
  JOURNAL   Unpublished (2003)
COMMENT     Other_ESTs: LeukoN2_8_B10.b2_A024
            Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
            the Human Genome Center, University of Tokyo Institute of Medical
            Science; tissue and RNA were prepared in the Department of Large
            Animal Medicine, University of Georgia; sequencing done in the
            Laboratory for Genomics and Bioinformatics, University of Georgia.
            Sequence ends have been trimmed to exclude vector and regions below
            Phred quality 16. Three-prime sequences are presented as their
            reverse complement and have been trimmed to exclude polyA.
            Seq primer: Sug5 (CTTCTGCTCTAAAAGCTGCG).
FEATURES             Location/Qualifiers
     source          1..661
                     /organism="Equus caballus"
                     /mol_type="mRNA"
                     /strain="thoroughbred"
                     /db_xref="taxon:9796"
                     /clone="LeukoN2_8_B10_A024"
                     /sex="male"
                     /tissue_type="blood"
                     /cell_type="leukocytes"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="Unstimulated peripheral blood leukocytes N2"
                     /note="Organ: circulatory system; Vector: pME18S-FLJ;
                     Site_1: XhoI; Site_2: XhoI; The library was prepared from
```

URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi.K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Chara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,Y., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

```
FEATURES       e mouse tissues.
     source    Location/Qualifiers
               1..616
               /organism="Mus musculus"
               /mol_type="mRNA"
               /db_xref="taxon:10090"
               /clone="9830006J10"
               /sex="male"
               /tissue_type="bone"
               /dev_stage="adult"
               /lab_host="DH10B"
               /clone_lib="RIKEN full-length enriched, adult male bone"
               /note="Site_1: SalI; Site_2: BamHI; cDNA library was
               prepared and sequenced in Mouse Genome Encyclopedia
               Project of Genome Exploration Research Group in Riken
               Genomic Sciences Center and Genome Science Laboratory in
               RIKEN. Division of Experimental Animal Research in Riken
               contributed to prepare mouse tissues. 1st strand cDNA was
               primed with a primer [5'
               GAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTVN 3'], cDNA was
               prepared by using trehalose thermo-activated reverse
               transcriptase and subsequently enriched for full-length by
               cap-trapper. cDNA went through one round of normalization
               to Rot = 10.0 and subtraction to Rot = 185.0. Second
               strand cDNA was prepared with the primer adapter of
               sequence [5' GAGAGAGAGATCCAGTTAATTAAATTAATCCCCCCCCCCCC
               3'].cDNA was cleaved with XhoI and BamHI. Vector: a
               modified pBluescript KS(+) after bulk excision from Lambda
               FLC I."
```

```
ORIGIN
Query Match    42.4%;    Score 493.8;  DB 10;  Length 616;
Best Local Similarity  98.6%;   Pred. No. 1e-94;
Matches 498;  Conservative 0;  Mismatches 7;  Indels 0;  Gaps 0;
```

```
Qy     1  TTTTTGTAAGAAAAATCTGAGGAAGAGATTCGGGATAGCGCTTTATTCAGGATGTTTCCT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    46  TTTTTGTAAGAAAAATCTGAGGAAGAGATTCGGGATAGCGCTTTATTCAGGATGTTTCCT  105
Qy    61  ATGAAATAGCATTCATCTGTGGGGAGGAGGAGGACTAAGGAAATCTGACATCTGTTGGTC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   106  ATGAAATAGCATTCATCTGTGGGGAGGAGGAGGACTAAGGAAATCTGACATCTGTTGGTC  165
Qy   121  ACTGGGACAGAATATGGGATATGTCACTGGTTTTCTTCCTCCTCAATTAAATATGTGAATC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   166  ACTGGGACAGAATATGGGATATGTCACTGGTTTTCTTCCTCCTCAATTAAATATGTGAATC  225
Qy   181  AAAAGAGAAAACTTTTTCAAACTAATGCATGGTCATGGTCAGGGAAGAAGAACAGCATCGA  240
```

JOURNAL   Unpublished (2003)
COMMENT   Contact: Smith TPL
           USDA, ARS, US Meat Animal Research Center
           PO Box 166, Clay Center, NE 68933-0166, USA
           Tel: 402 762 4366
           Fax: 402 762 4390
           Email: smith@email.marc.usda.gov
           Single pass sequencing. Bases called with phred v0.020425.c and
           trimmed with the aid of the trim_alt option. Vector identified with
           cross_match v0.990329.
           Plate: SRG8012 row:O column:18
           Seq primer: GTAATACGACTCACTATAGGG.

FEATURES   Location/Qualifiers
      source   1..723
                    /organism="Sus scrofa"
                    /mol_type="mRNA"
                    /db_xref="taxon:9823"
                    /tissue_type="pooled"
                    /lab_host="DH10B"
                    /clone_lib="MARC 3PIG"
                    /note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
                    Library made with RNA pooled from multiple tissues
                    including brain, liver, muscle, placenta/endometrium,
                    ovary, testes, and bone marrow."

ORIGIN

Query Match   40.3%;   Score 469.6; DB 14;   Length 723;
Best Local Similarity 79.0%;   Pred. No. 1.4e-89;
Matches   559;   Conservative   0;   Mismatches 149;   Indels   0;   Gaps   0;

Qy   86 AGAGAAGGACTAAGGAAATCTGACATCTGTTGGTCACTGGGACAGAATATGGATATGTCA 145
            ||||| ||||| || ||||||||||||| |||| ||||| || |||||| || ||| ||
Db   16 AGTGGAAGATGTAATAAATATTACAACTCCTTCACTTTATAGAGAGAAGATGGATACACCC 75

Qy   146 CTGGTTTTCTTCTCTCAATTAAATATGTGAATCAAAAGAGAGAAACTTTTTCAAACTA 205
            ||||||||||||||||| |||||||||| |||||||||||||||||||||||||||||||
Db   76 CTGCTTTTCTTTCTCTCAATTAAATATGTGAACCAAAAGAGAGAAAACTTTTTCAAGTTA 135

Qy   206 ATGCATGGTCAGGGAAGAGAAGAGAAGAAGAGAAGCAAGAGAGCGAGGCCAAAATCAGAGG 265
            ||||||||||||||||||||||||| ||||||||||||||||| ||| ||||||||||||
Db   136 ATTCATGGTTCAGGGAAGAGAAGAAGAGAAGAGAAGCAAGAGAGCCAGAGGACTCAAGCCAGT 195

Qy   266 AGGAATAGACTAAGTCTTCTCCTGAAGAAACAGAGGCCTGACTTCCATGGGACAGAATTAT 325
            |||| |||||| |||||||||||||| |||||||||||||||||||||| |||||||| |
Db   196 AGAAACAGGCTAAGTCTTCTCTCGTCGCAGAAAACCTGAGTTTCACGATGTACTCCCTCCAGT 255

Qy   326 AGATCTGCCCTTGGCCAAAGCAGAGTCTCTCCTGAAGAAACAGAGGCAGTCAGTGAAATGGGCT 385
            ||||| |||||||||||||||||||||| ||||||||||||||||||||||| ||||||||
Db   256 AGTTCTGAACACCTGCCGCCAAAAAAAGAATCAATCAAGAGAGCAAATGATATATGGGGT 315

Qy   386 GAATCATTGACAAATTGCTCTCTCATGAGATGATGCCCCAAA 445
            ||||||||||||||||||| |||||||| |||||| ||||||
Db   316 GAATCATTGACAAACTGCTTTCCCAGGAAAGATGATTGAATTTGGTCGCCGTGTGGAAGACTTCAAGAAA 375

Qy   446 AAAACTGAATTCAGTGAGGAGAACATTGAATTTGGTCGCCGTGTGGAAGACTTCAAGAAA 505
            ||||||||||||||||||||| || || |||||||||||||||||||||| |||||||||
Db   376 AAGACCGAATTCAGTGATGATAACATTGCATTTTGGATAGCCTGTGGAAGATTACAAGAAA 435

Qy   506 TGCAAGGAACCTCAACAAATCATCCTAAAAGGCAAAGCAATCTATGAGAAATTCATTCAG 565
            || |||||||||| ||||| || |||||||||||||| |||| || ||||||| |||||
Db   436 AGCAAGGACCCTCAACAAATAATTCGGAAAGCGAAAGCAATATATGAGACGTTTATACAG 495

Qy   566 AATGATGCCCCAAAGAGAGTTAACATTGATTTCATACTAAGAAGTAATTGCTAAGAGC 625
            |||||| || ||||||||| ||||||||||||| |||||||||||||| || ||||||
Db   496 ACCGATGCTCCACAAGATGTTAACCTTGATTTTCACCCAAAGAAAATCCTCGCCAAGATC 555

Qy   626 ATCGCCCAGCCCACTCCCACATCTCCAACAGTTTTGATACGCGCACAAAGCAGAGTGTACCAGCTCATG 685
            ||||| ||| ||||| |||||| ||||||||||||||| |||| |||| |||||||| |||||||||||
Db   556 ATCACCGAGCCTACCCTCCACAGCTTTGATCGTCGCGCATAGCATGTATCATCCATG 615

Qy   686 GAACATGACAGTTACAACGCTTTTGAAATCTGACGTTATAAACGCTTTTGAAATCTGAGACCTACTTACATTTGATAGAAGGA 745
            ||||| ||||||| || |||||||||||||||| ||||||||||||||| |||||||||||||||||| |||||||||||||
Db   616 GAACAAGACAGTTACAACACCTGCTTTTCTGAAATCCGACATCATTATTGACTTAGACTTGATAGAAAGA 675

RESULT 11
CF181629
LOCUS   CF181629   723 bp   mRNA   linear   EST 28-JUL-2003
DEFINITION   81847B MARC 3PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   CF181629
VERSION   CF181629.1   GI:33293405
KEYWORDS   EST.
SOURCE   Sus scrofa (pig)
ORGANISM   Sus scrofa
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1 (bases 1 to 723)
AUTHORS   Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Fox,J.,
             Wise,T.A., Nonneman,D.J., Wray,J.E. and Keele,J.W.
TITLE   A second set of porcine ESTs from a pooled-tissue normalized
             library

ORIGIN

Query Match   40.9%;   Score 476; DB 14;   Length 661;
Best Local Similarity 83.1%;   Pred. No. 6.2e-91;
Matches   542;   Conservative   0;   Mismatches 110;   Indels   0;   Gaps   0;

Qy   101 AAAATCGACATCTGTTGGTCACTGGGACAGAATATGGATATGTCACTGGTTTTCTTCTCT 160
            |||| ||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   10 AAATTAGACGTCTCTTCATTTGAGAAAGAAGAGTGGAACATCATTGGTTTTCTTTTTCT 69

Qy   161 CAATTAAATATGTGAATCAAAAGAGAGAAACTTTTTCAAACTAATAATGCATGGTCAGGGG 220
            |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   70 CAATTAAATATGTGCGAATCAAAAGAAAAAAAACTTTTTCAAGTTAATACACAGGTTAAGGA 129

Qy   221 AAGGAGAAACAAGCATCGAGGCCAAAATCAGAGCGAAAGAAAAAAAGGAATAGACTAAGT 280
            |||| |||| ||||||||||||||||||||||||||||||||||||||| ||||| ||||
Db   130 AAAGAGAAACAAGCAAAGAGAAGCCAAACTCAGAGCTAAGGAAAAAAGAAACAGGTTAAGT 189

Qy   281 CTTCCTACAGAGAGCCTGACTTCCATGGAGAGACTCAAGCCAGTAGACTCGCCCTCTG 340
            |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   190 CTGCTCTTGCGAGAAACCTGAATTTCATGAAGAGATCCTCCGGTAGAGATCGGGCACTTA 249

Qy   341 GCCAAAGAAACAAGAGTCCTCTCCCTGAAGAAGCCAGTGAAATGGGCCGTGAATCATTTGACAAA 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   250 GCCAAAGAAACAGCAGAATCTCCCCTGCAGAGACCGTGATATGGGTGAATCATTTGACAAA 309

Qy   401 TTGCTCCTCATAGATGGAGTCGGATCGGACTGGAGACTTTTACCAGATTCCTAAAACTGAATTCAGT 460
            |||| |||||||||||||||||||||||||||| |||| ||||||||||||||||||||||
Db   310 CTGCTCTCCATAAAGGTCGACTCGGAGACTTTTTACCAGATTCCTAAAACCGAATTCAGT 369

Qy   461 GAGGAGAACATTGAATTTGGGTCGCCTCGTGAAGAATTCAAGGAAATGCAAGGAACCTCAA 520
            ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   370 GAGGAAAACATTGAATTTGGGTAGCCTCGTGAAGAATTTCAAGGAAAAGCAAAGAACCCTCAA 429

Qy   521 CAAATCATCCTAAAAGCAAAGCAATCTATGAGAAATTCATTGAGAAATTGACGAATGCCCCCAAA 580
            |||| ||||| |||||||||||||||||||||||||| |||||||| |||||||||||||||||
Db   430 CAAATTATCCTTAAAGCAAAAGTAATATATGAGACAATATGCTAAGAGCATGCGTCCGCAA 489

Qy   581 GAGGTAACATTGATTTTCATACTAAAGAGTAATTGCTAAGAGAGCATCGACGCCCAGCCCACT 640
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   490 GAGGTTAAACCTTGATTTCATTTCCACACCAAAGAAAGAGCAGTAGCACCAAACCCACT 549

Qy   641 CTTCCACAGTTTTGATACGGCGACACAAAGCAGAGTGTACCAGCTCATGGGAACATGACAGTTAT 700
            |||||||||||| |||||||| ||||| |||||||||||||||| ||||||||||||| |||
Db   550 CTTCCACAGTTTTGATGCGTCGCACAAAGCAGTATCAGCGCTTATGGGAACAAGACAGTTAC 609

Qy   701 AAACGGCTTTTTTTGAAATCTGAGAACCTATCTATTTAGACTTTGACAGAAGGAAGGAGACCTC 752
            ||| ||||||| |||||||||| ||| |||||||||||||||||||||||| ||||||||||
Db   610 ACACGTTTTTCTGAAATCCGACATCATCATTATTGACTTTGACAGAAGGAAGGAGACCTC 661

RESULT 11
CF181629
LOCUS   CF181629   723 bp   mRNA   linear   EST 28-JUL-2003
DEFINITION   81847B MARC 3PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   CF181629
VERSION   CF181629.1   GI:33293405
KEYWORDS   EST.
SOURCE   Sus scrofa (pig)
ORGANISM   Sus scrofa
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1 (bases 1 to 723)
AUTHORS   Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Fox,J.,
             Wise,T.A., Nonneman,D.J., Wray,J.E. and Keele,J.W.
TITLE   A second set of porcine ESTs from a pooled-tissue normalized
             library

Qy　746 AGACCTCAGAGACCAACAAACCTTAGGAGAGGATCACGATCATTTACT 793
Db　676 AGACCTCGGAGGACGAACAAATCTGAGAAGGCGGATCACGTTCCTTTACT 723

```
RESULT 12
CD468328
LOCUS       CD468328               764 bp    mRNA    linear   EST 04-JUN-2003
DEFINITION  LeukoS3_2_G10.g1_A025 stimulated peripheral blood leukocytes S3
            Equus caballus cDNA clone LeukoS3_2_G10_A025 5', mRNA sequence.
ACCESSION   CD468328
VERSION     CD468328.1  GI:31389596
KEYWORDS    EST.
SOURCE      Equus caballus (horse)
  ORGANISM  Equus caballus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE   1  (bases 1 to 764)
  AUTHORS   Vandenplas,M., Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S.,
            Moore,J.N., Liang,C., Sun,F., Sullivan,R., Shah,M. and Pratt,L.H.
  TITLE     An EST database from equine (Equus caballus) stimulated peripheral
            blood leukocytes
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
            the Human Genome Center, University of Tokyo Institute of Medical
            Science; tissue and RNA were prepared in the Department of Large
            Animal Medicine, University of Georgia; sequencing done in the
            Laboratory for Genomics and Bioinformatics, University of Georgia.
            Sequence ends have been trimmed to exclude vector and regions below
            Phred quality 16. Three-prime sequences are presented as their
            reverse complement and have been trimmed to exclude polyA.
            Seq primer: Sug5 (CTTCTGCTCTAAAAGCTGCG).
FEATURES             Location/Qualifiers
     source          1..764
                     /organism="Equus caballus"
                     /mol_type="mRNA"
                     /strain="Tennessee walking horse"
                     /db_xref="taxon:9796"
                     /clone="LeukoS3_2_G10_A025"
                     /sex="female"
                     /tissue_type="blood"
                     /cell_type="leukocytes"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="Stimulated peripheral blood leukocytes S3"
                     /note="Organ: circulatory system; Vector: pME18S-FL3;
                     Site 1: XhoI; Site 2: XhoI; The library was prepared from
                     polyA+ RNA from equine peripheral blood leukocytes
                     isolated from a healthy adult horse. The leukocytes were
                     stimulated for 4 hr with 10 ng/ml E. coli O55:B5 LPS.
                     Double-stranded cDNA was cloned unidirectionally into
                     different DraIII sites of the pME18S-FL3 vector (5-prime
                     DraIII site is CACTGTG, 3-prime DraIII site is
                     CACCATGTG). XhoI excises the cDNA insert."
ORIGIN
```

Query Match　　　37.7%;　　Score 438.6;　　DB 14;　　Length 764;
Best Local Similarity　82.7%;　　Pred. No. 5.5e-83;
Matches 501;　Conservative　0;　Mismatches 105;　Indels 0;　Gaps 0;

Qy　 87 GAGAGGACTAAGGAAATCTGACANTCTGTTCGTCACTGGGACAGAATATGGATATGTCAC 146
Db　159 GTGAAGGATGTAATAAATTAGACGCTCTCTCATTTTAGAGAAAGAGATGGAAACATCAT 218
Qy　147 TGGTTTCTTTCTCAATTAAATATGTGTGAATCAAAAGAAAGAGAAACTTTTTCAAACTAA 206

Qy　219 TGGTTTCTTTCTCAATTAAATATGTCGAATCAAAAGAAAAAACTTTTTCAAGTTAA 278
Db
Qy　207 TGCATGGGTCAGGGAAAGAAAGAAACAGCATCGAGGCCAAATCAGAGCGAAAGAAAAA 266
Db　279 TACATGGTTCAGGAAAAGAAAGAAGAAAACAGCAAAGAAGCCAAATCAGAGCTAAGGAAAAA 338

Qy　267 GGAATAGACTAAGTCTTCTCCTACAGAGGCCTGCACTTCCATGGAGAGACTCAAGCCAGTA 326
Db　339 GAAACAGGTTAAGTCTGCTTCTTGCAGAAACCTGAATTTCATGGAAGAGACCCACTCCGGTA 398

Qy　327 GATCTGCCCTCTTGGCAAAGAAACAACAGAGTCTCCTCTGAAGAATCATTGAGATGGGCTG 386
Db　399 GATCCGGCACTTAGCCAAAGAAAACAGAAGAATCTCCCCTGCAGAAGCCGTGATATGGGGTG 458

Qy　387 AATCATTTGACAAATTGCTTCTCATAGAGATGGAGTGGATGCTTTTACCAGATTTCTTA 446
Db　459 AATCATTTGACAAACTGCTCTCCCATAAAGGTGGACTGGAGACTTTTACCAGATTTCTTA 518

Qy　447 AAACTGAATTCAGTGAGGAGAACATTGAATTTTGGGTGCGCCTGTGAGGACTTCAAGAAAT 506
Db　519 AAACCGAATTCAGTGAGGAAAACATTGAATTTTGGGTAGCCTGTGAAGATTTCAAGAAAA 578

Qy　507 GCAAGGAACCTCAACAAATCATCCTAAAAGCAAGGCAATCTCATGAGAAATTCATTCAGA 566
Db　579 GCAAGGACCCTCAACAAATTATCCTTAAAGCAAAAGCAATAATGAGAAATTTATACAGA 638

Qy　567 ATGATGCCCCAAAGAGGTTAACATTGATTTTCATACATAAAGAAGTAATTGCTAAGAGCA 626
Db　639 ACGATGCTCCGCAAGAGGTTAACCTTGATTTCCACCACCAAAGAAATCATTGCCAAGAGCA 698

Qy　627 TCGCCCAGCCACTCTCCACAGTTTTGGATACGGCACAAAGCAGGAGTGTACCAGCTCATGG 686
Db　699 TCACCAAACCCACTCTCCACAGNTTTGATGCTGCACAAAGCAGAGTATATCAGCTTATGG 758

Qy　687 AACATG 692
Db
Qy　759 AACAAG 764

```
RESULT 13
BG564257
LOCUS       BG564257               810 bp    mRNA    linear   EST 10-APR-2001
DEFINITION  602586054F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4714909 5',
            mRNA sequence.
ACCESSION   BG564257
VERSION     BG564257.1  GI:13571909
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 810)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
            cDNA Library Preparation: CLONETECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1560 row: 1  column: 14
            High quality sequence stop: 772.
FEATURES             Location/Qualifiers
     source          1..810
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4714909"
                     /lab_host="DH10B (T1 phage-resistant)"
```

```
DEFINITION    AV714060 DCB Homo sapiens cDNA clone DCBBIF02 5', mRNA sequence.
ACCESSION     AV714060
VERSION       AV714060.1  GI:10795577
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 705)
AUTHORS       Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
              Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
              Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
              Lu,G., Cheng,Z. and Han,Z.
TITLE         Homo sapiens cDNA DCB clones
JOURNAL       Unpublished (2000)
COMMENT       Contact: Zeguang Han
              Chinese National Human Genome Center at Shanghai
              351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
              201203, P. R. China
              Tel: 86-21-50801919(ex.45)
              Fax: 86-21-50801922
              Email: hanzg@chgc.sh.cn
              This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1..705
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DCBBIF02"
                     /cell_type="dendritic cells"
                     /dev_stage="mature"
                     /lab_host="BM25.8"
                     /clone_lib="DCB"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
ORIGIN
```

```
Query Match       33.7%;    Score 392.2;   DB 9;   Length 705;
Best Local Similarity   80.1%;   Pred. No. 4e-73;
Matches 486;   Conservative   0;   Mismatches 118;   Indels   3;   Gaps   2;

Qy      4 TTGTAAGAAAAATCTGAGGAAGAGATTCGGGATAGCGCTTATTC-AGGATGTTTCCTAT   62
          |||| ||||||| ||||||||| ||||| |||||||||||| |||| ||| |||||||||
Db     41 TTTTTTGTAAACATTACTGTAAGAGTTGTGATAACTTTTATTCTACTACTATATGTAT  100

Qy     63 GAAATAGCATTCATCTGTGGGAGAGAAGAAGACTAAGGAAATCTGACATCTGTTGGTCAC  122
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    101 GGAATAGTATTAATAAATGAACTAGGAGGGATGTAATAAATTAGACATCTTCATT--  158

Qy    123 TCGGACAGAATATGGATATGTCACTGTTTCTTCTCAATTAAATATGTGAATCAA  182
          | | |||||||||| ||||||||||||||| ||||| ||||||||||||||||||
Db    159 TTAGAGAAGAAGATGGAAACACCATTGCTTTTCTTTTTCTCAAATAAATATGTGAATCAA  218

Qy    183 AAGAGAGAAAACTTTTTTCAAACTAATGCATGGGTCAGGGAAAGAAGAGAAAACAAGCATCGAGG  242
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    219 AAGAAAAAACTTTTTTCAAGTAATCACGTTTCAGGAAAAGAAGAGAAAACAAGCAAAGAAG  278

Qy    243 CCAAAATCAGAGACGAAGAAGAAAAAGAAATAGAAGAAAGAGAAATAGACTAAGTCTTCTCCTACAGAGGGCCGACT  302
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    279 CCAAAATCAGAGACGAAGCTAAGGAAAAGAAATAGAACTAAGTCTTCTTGTGCAAACAAGAAGTCTGAGT  338

Qy    303 TCCATGGAGAGACTCAAGCCAGTAGATCTGCCCTCTGGCCAAAGAAACAAGAGTCTCTC  362
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    339 TTCATGAAGAAGACACCCGCTCCAGTAGATCTCGGGCACTCGGGCACTCGGGCAAAGAAACAAGAGTCTCCC  398

Qy    363 CTGAAGAAGCAGTAGAATGGAAATGCTCTCTCATAGATGGAG  422
          ||||||||||||||||||||||||||||||||||||||||||
Db    399 CTGAAGAGGCAGTGAAATGGGGTGAATCATTTCCCATAGAGATGGAC  458

Qy    423 TGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTAGGAGAAAATATTGAATTTTGGG  482
          |||||||||||||||| ||||||| ||||||||||||| |||||| ||||| ||||||||
Db    459 TAGAGGCTTTTTTACCAGATTTCTTAAAACTGAATTCAGTAGTGAGAAAAATATTCATCCTAAAGACAAAG  518

Qy    483 TCGCCTGTGAAGACTTCAAGAAATGCAAGAAATCCTCAACAAATCATCCTAAAGACAAAG  542
```

```
RESULT 14
AV714060
LOCUS       AV714060                705 bp    mRNA    linear    EST 11-OCT-2000
```

```
Db   266 CCAAAATCAGAGCTAAGGAAAAAAGAAATAGACTAAGTCTTCTTGTGTGCAGAAACCTGAGT 325

Db   519 |||||||||||||||||||||||||||||||||||||||||||||||||| ||| |||||||
     TAGCCTGTGAAGATTCAAGAAAAAGCAAGGGACCTCAACAAATTCACCTTAAAGCAAAAG 578

Qy   303 TCCATGGAGAGACTCAGCCAGTAGACTCGCCCTCTTGGCCAAAGAAACAAGAGTCTCTC 362
         |||| ||||||||||||||||||||||||   ||||||||||||||||||||||||
Db   326 TTCATGAAGACACACCCGCTCCAGTAGATCTGGGCACTTGGCCAAAGAAACAAGAGTCTCCC 385

Qy   543 CAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTTCATA 602
         ||| ||||||||||||| |||| |||||||||||| |||| ||||||||||||||||||
Db   363 CTGAAGAAGCAGTGAAATGGGCTGAATCATTGACAATTGCTCTCTCATAGAGATGGAG 422

Qy   579 CAATATATAGAGAAATTTATACAGACTGATGCCCCAAAAGAGGTTAACCTTGATTTTCACA 638
         ||||| ||||||||||| |||| |||||||||||| ||||||||||||||||||||||
Db   386 CTGAAGAGGCAGTGAAATGGGGTGAATCATTGACAAACTGCTTTCCCATAGAGATGGAC 445

Qy   603 CTAAAGA 609
         ||| ||
Db   423 TGGATGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAGAGAACATTGAATTTTGGG 482

Db   639 CTGCACA 645
         ||| |||
Db   446 TAGAGGCTTTTACCAGATTTCTTAAAACTGAATTCAGTGAAGAAAATATTGAATTCTGGA 505

Qy   483 TCGCCCTGTGAAGACTTCAAGAAATGCAAGGAACCTCAACAAATCATCCTAAAAGCAAAGG 542
         |||||||||||||||||||||| ||||||||| ||| |||||||||||||||||||||
Db   506 TAGCCTGTGAAGATTTCAAGAAGATTAAGAAGCAAGGGACCTCAACAAATTCACCTTAAAGCAAAAG 565

Qy   543 CAATCTATGAGAAATTCATTCAGAATGATGCCCCAAAGAGGTTAACATTGATTTTCATA 602
         ||| ||||||||||||| |||| |||||||||||| ||||||||||||||||||| ||||
Db   566 CAATATATGAGAAATTTATACAGACTGATGCCCCAAAAGAGGTTAACCTTGATTGTCACA 625

Qy   603 C 603
         |
Db   626 C 626
```

Search completed: August 20, 2004, 12:31:53
Job time : 3160.02 secs

```
RESULT 15
AV714044
LOCUS       AV714044                626 bp    mRNA    linear    EST 11-OCT-2000
DEFINITION  AV714044 DCB Homo sapiens cDNA clone DCBCFH06 5', mRNA sequence.
ACCESSION   AV714044
VERSION     AV714044.1  GI:10795561
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 626)
AUTHORS     Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Cheng,Z. and Han,Z.
TITLE       Homo sapiens cDNA DCB clones
JOURNAL     Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1..626
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DCBCFH06"
                     /cell_type="dendritic cells"
                     /dev_stage="mature"
                     /lab_host="BM25.8"
                     /clone_lib="DCB"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
ORIGIN
Query Match        33.3%;    Score 387.8;  DB 9;   Length 626;
Best Local Similarity  80.0%;    Pred. No. 3.5e-72;
Matches 481;  Conservative    0;  Mismatches 117;  Indels    3;  Gaps    2;

Qy     4 TTGTAAGAAAAATCTGAGGAAAGATTCGGGATAGCGGCTTATTC-AGGATGTTTTCCTAT 62
         || ||| |||||  |||| |||||||| ||| || | || ||| |||||||| |||||
Db    28 TTTTTTGTAAACATTACTGTAAGAGTTGTGATAACTTTTTATTCTACTATGTATATGTAT 87

Qy    63 GAAATAGCATTCATCTGTGGGAGGAGAGGAGGACTAAGGAAATCTGACATCTGTTGGTCAC 122
         ||||||| |||||| | |||||||||| ||||||||||||||| ||||| |||||||||||
Db    88 GGAATAGTATTAAATGAACTAGGGAAGGATGTAATAAATTAGACATCTCTTCATT-- 145

Qy   123 TGGACAGAGATATGATATGTCACTGGTTTCTCTCAATTAAATATGTGAATCAA 182
         ||||| ||| |||| ||| |  |||||||||| |||| |||||||||||||||
Db   146 TTAGAGAGAGATGGAAACAACATTGCTTTGCTTTTCTTTCTCAAATAAATAAATGTGAATCAA 205

Qy   183 AAGAGAAAACTTTTTCAAACTAATGCATGCATGGGTCAGGGAAGAAATCGAGGAAGCATCAGAGG 242
         |||| ||| ||||||||||||| || ||||| ||| ||||||| |||||| |||| ||||||
Db   206 AAGAAAAAACTTTTTTCAAGTTAATACACATGGTTCAAGTTAATACACATGGTTCAGGAAAGAAAAGAAGCAAGCAAGAAG 265

Qy   243 CCAAAATCAGAGCGAAAGAAAGAAAAAGGAATAGACTAAGTCTTCTTCCTACAGAGAGGCCTGACT 302
         ||||||||||||| ||| ||| ||| ||| |  ||||||| ||||| |
```

```
              GenCore version 5.1.6
         Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      August 20, 2004, 12:36:30 ; Search time 79 Seconds
                               (without alignments)
                     840.489 Million cell updates/sec

Title:       US-09-894-749-4
Perfect score: 1214
Sequence:    1 MDMSLVFFSQLNMCESKEKT........SRSFTYNDFQDVKSDVAIWL 235

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:    1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:     1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   A_Geneseq_29Jan04:*
           1: geneseqp1980s:*
           2: geneseqp1990s:*
           3: geneseqp2000s:*
           4: geneseqp2001s:*
           5: geneseqp2002s:*
           6: geneseqp2003as:*
           7: geneseqp2003bs:*
           8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                         SUMMARIES
```

Result No.	Score	% Query Match	Match Length	DB	ID	Description				
1	1214	100.0	235	3	AAY97154	Aay97154 Murine re		26	388	32.0	152	4	AAG66509	Aag66509 RGS prote
2	1214	100.0	235	5	AAE25830	Aae25830 Murine re	27	383	31.5	519	4	AAB90786	Aab90786 Human she
3	1032	85.0	235	3	AAY97153	Aay97153 Human reg	28	383	31.5	519	6	ABB99466	Abb99466 Amino aci
4	1032	85.0	235	5	AAE25829	Aae25829 Human reg	29	383	31.5	917	6	ADA13333	Ada13333 Human int
5	1032	85.0	235	5	AAU10749	Aau10749 Human RGS	30	380	31.3	284	6	ADA54082	Ada54082 Human pro
6	1032	85.0	235	5	ABP64801	Abp64801 Human pro	31	375.5	30.9	923	2	ABU54617	Abu54617 Human NOV
7	994	81.9	227	5	ABP64054	Abp64054 Human ORF	32	374	30.8	118	2	AAY41001	Aay41001 Mouse RGS
8	429.5	35.4	211	5	ABP96779	Abp96779 Human COP	33	369	30.4	116	2	AAY41020	Aay41020 RGS2 prot
9	429.5	35.4	211	6	ABP98020	Abp98020 Amino aci	34	364	30.0	202	2	AAR89895	Aar89895 p53 respo
10	429.5	35.4	211	7	ADD14019	Add14019 Human src	35	364	30.0	202	6	ABR92146	Abr92146 Human cer
11	429.5	35.4	220	3	AAB56929	Aab56929 Human pro	36	363	29.9	202	2	AAW08133	Aaw08133 Human cyt
12	427.5	35.2	181	2	AAW62075	Aaw62075 Human reg	37	363	29.9	202	2	AAW59294	Aaw59294 Human RAT
13	427.5	35.2	181	5	AAY53931	Aay53931 A human r	38	363	29.9	202	3	AAY87952	Aay87952 Human CR1
14	427.5	35.2	190	5	ABP62144	Abp62144 Human sec	39	362	29.8	81	5	AAU10747	Aau10747 Amino aci
15	412.5	34.0	181	5	ABB83788	Abb83788 Rat disea	40	360	29.7	128	7	ABR82620	Abr82620 Rat RGS4
16	410	33.8	198	4	ABB62187	Abb62187 Human RGS	41	360	29.7	167	5	AAU11043	Aau11043 Rat RGS4
17	407.5	33.6	180	5	ABP69254	Abp69254 Human pol	42	356	29.3	201	2	AAW59293	Aaw59293 Mouse RAT
18	407.5	33.6	180	5	ABB04999	Abb04999 Human reg	43	353.5	29.1	249	6	ABP98019	Abp98019 Amino aci
19	397.5	32.7	205	7	ADE62388	Ade62388 Human RGS	44	350	28.8	111	6	ABP60081	Abp60081 Human RGS
20	397.5	32.7	205	7	AAU78976	Aau78976 Human Pro	45	347.5	28.6	196	4	AAB76863	Aab76863 Human lun
21	397.5	32.7	205	6	ADE62392	Ade62392 Human Pro							
22	396	32.6	930	6	ABB99465	Abb99465 Amino aci							
23	395.5	32.6	205	7	ADE62386	Ade62386 Rat Prote							
24	395.5	32.6	205	7	ADE62390	Ade62390 Rat Prote							
25	388	32.0	135	6	ABP98021	Abp98021 Amino aci							

```
                              ALIGNMENTS

RESULT 1
AAY97154
ID   AAY97154 standard; protein; 235 AA.
XX
AC   AAY97154;
XX
DT   04-DEC-2000  (first entry)
XX
DE   Murine regulator of G-protein signaling protein.
XX
KW   RGS; regulators of G-protein signaling; GTPase activating protein; GAP;
KW   G-alpha protein; cell adhesion; chemotaxis; vulnerary; immunosuppressor;
KW   anti-rheumatic; anti-arthritic; anti-diabetic; anti-inflamatory;
KW   cytostatic; hepatotropic; anti-anaemic; modulator; gene therapy.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          82..201
FT                   /label= RGS_domain
FT   Misc-difference 83
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 90
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 100
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 103
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 104
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 107
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 109
FT                   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 111
FT                   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 112
FT                   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 115
FT                   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 116
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 138
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 139
FT                   /note= "forms part of hydrophobic core"
FT   Misc-difference 142
FT                   /note= "forms part of hydrophobic core"
```

```
Qy      121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180

Qy      181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235


RESULT 2
AAE25830
ID  AAE25830 standard; protein; 235 AA.
XX
AC  AAE25830;
XX
DT  15-NOV-2002  (first entry)
XX
DE  Murine regulator of G-protein signalling (RGS) protein, m1975.
XX
KW  Murine; screening; RGS; regulator of G-protein signalling; pancreatitis;
KW  inflammatory disorder; allergy; Grave's disease; arthritis; sinusitis;
KW  respiratory disorder; asthma; pneumonia; therapy; immune disorder;
KW  haematological disorder; haematopoiesis; platelet-associated disorder;
KW  thrombocytopaenia; migration; invasive disorder; leukaemia; anaemia;
KW  erythrocyte-associated disorder; hepatitis; m1975.
XX
OS  Mus sp.
XX
FH  Key             Location/Qualifiers
FH
FT  Domain          82..201
FT                  /note= "RGS domain"
XX
PN  US6410240-B1.
XX
PD  25-JUN-2002.
XX
PF  04-FEB-2000;  2000US-00498959.
XX
PR  04-FEB-1999;  99US-00244314.
XX
PA  (MILL-) MILLENIUM PHARM INC.
XX
PI  Hodge MR, Yowe D;
XX
DR  WPI; 2002-588886/63.
DR  N-PSDB; AAD42498.
XX
PT  screening assays for identifying agent that binds to human or mouse RGS
PT  (regulators of G-protein signalling) protein or its variant, or RGS
PT  protein encoded by homologous DNA sequences, or to host cell expressing
PT  the proteins.
XX
PS  Claim 1; Col 57-60; 42pp; English.
XX
CC  The invention relates to screening assays for identifying an agent that
CC  binds to: human or mouse RGS (regulators of G-protein signalling) protein
CC  or its variant, or RGS protein encoded by homologous DNA sequences; or a
CC  host cell expressing the RGS protein or its variant, or RGS protein
CC  encoded by homologous DNA sequences. The invention is useful for
CC  identifying an agent that binds to human or mouse RGS protein or its
CC  variant, or RGS protein encoded by homologous DNA sequences, or a host
CC  cell expressing the RGS protein or its variant, or RGS protein encoded by
CC  homologous DNA sequences. The agents identified using the invention are
CC  useful for modulating the activity of RGS proteins and thus useful for
CC  treating immune and inflammatory disorders (Grave's disease, allergy,
CC  arthritis), respiratory disorders (asthma, pneumonia, sinusitis),
CC  haematological disorders (haematopoiesis, migration), platelet-associated
CC  disorders (thrombocytopaenia), invasive disorders (leukaemia)
CC  erythrocyte-associated disorders (anaemia), pancreatitis, hepatitis etc.
CC  The present sequence is murine RGS protein, m1975
XX
SQ  Sequence 235 AA;
```

```
FT  Misc-difference 143
FT                  /note= "forms part of hydrophobic core"
FT  Misc-difference 151
FT                  /note= "forms part of hydrophobic core"
FT  Misc-difference 152
FT                  /note= "Makes direct contact with G-alpha-i"
FT  Misc-difference 154
FT                  /note= "Makes direct contact with G-alpha-i"
FT  Misc-difference 183
FT                  /note= "Makes direct contact with G-alpha-i"
FT  Misc-difference 184
FT                  /note= "forms part of hydrophobic core"
FT  Misc-difference 187
FT                  /note= "Makes direct contact with G-alpha-i"
FT  Misc-difference 188
FT                  /note= "Makes direct contact with G-alpha-i"
FT  Misc-difference 189
FT                  /note= "forms part of hydrophobic core"
FT  Misc-difference 191
FT                  /note= "Makes direct contact with G-alpha-i"
FT  Misc-difference 192
FT                  /note= "forms part of hydrophobic core"
FT  Misc-difference 193
FT                  /note= "forms part of hydrophobic core"
FT  Misc-difference 198
FT                  /note= "forms part of hydrophobic core"
XX
PN  WO200046236-A2.
XX
PD  10-AUG-2000.
XX
PF  04-FEB-2000;  2000WO-US002977.
XX
PR  04-FEB-1999;  99US-00244314.
XX
PA  (MILL-) MILLENNIUM PHARM INC.
XX
PI  Hodge MR, Yowe D;
XX
DR  WPI; 2000-532893/48.
DR  N-PSDB; AAA52090.
XX
PT  Novel regulator of G-protein signaling nucleic acids and polypeptides,
PT  useful as diagnostic and investigative tools and to treat G-protein
PT  signaling disorders.
XX
PS  Claim 8; Page 104; 105pp; English.
XX
CC  The RGS (regulators of G-protein signaling) protein genes, clones
CC  AAH16395 and m1975, were identified in human and murine spleen cDNA
CC  libraries, respectively. Both proteins have unique N- and C-terminal
CC  sequences. The C-terminal location of the RGS domain is consistent with
CC  RGSs known to act as GTPase activating proteins (GAPs) for G-alpha
CC  proteins. G-alpha-i linked receptors support rapid adhesion and directed
CC  migration of leukocytes and other cell types. The novel RGS proteins may
CC  be used to modulate cell adhesion and chemotaxis, e.g. for aiding wound
CC  repair. The RGS proteins, related cDNAs and anti-RGS antibodies are
CC  useful for modulation, diagnosis and treatment of immune and respiratory
CC  disorders
XX
SQ  Sequence 235 AA;
```

```
Query Match             100.0%;  Score 1214;  DB 3;  Length 235;
Best Local Similarity   100.0%;  Pred. No. 1e-107;
Matches 235;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60

Qy      61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
```

```
Query Match         100.0%;     Score 1214;  DB 5;  Length 235;
Best Local Similarity  100.0%;  Pred. No. 1e-107;
Matches   235;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60

Qy   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120

Qy  121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180

Qy  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
```

RESULT 3
AAY97153
ID   AAY97153 standard; protein; 235 AA.
XX
AC   AAY97153;
XX
DT   04-DEC-2000 (first entry)
XX
DE   Human regulator of G-protein signaling protein.
XX
KW   RGS; regulators of G-protein signaling; GTPase activating protein; GAP;
KW   G-alpha protein; cell adhesion; chemotaxis; vulnerary; immunosuppressor;
KW   anti-rheumatic; anti-arthritic; anti-diabetic; anti-inflamatory;
KW   cytostatic; hepatotropic; anti-anaemic; modulator; gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   Domain   82..201
FT   /label= RGS_domain
FT   Misc-difference 83
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 90
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 100
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 103
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 104
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 107
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 109
FT   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 111
FT   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 112
FT   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 115
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 116
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 138
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 139
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 142
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 143
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 151
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 152
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 154
FT   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 183
FT   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 184
FT   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 187
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 188
FT   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 189
FT   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 191
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 192
FT   /note= "Makes direct contact with G-alpha-i"
FT   Misc-difference 193
FT   /note= "forms part of hydrophobic core"
FT   Misc-difference 198
FT   /note= "forms part of hydrophobic core"
XX
PN   WO200046236-A2.
XX
PD   10-AUG-2000.
XX
PF   04-FEB-2000; 2000WO-US002977.
XX
PR   04-FEB-1999; 99US-0244314.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Hodge MR, Yowe D;
XX
DR   WPI; 2000-532893/48.
DR   N-PSDB; AAA52089.
XX
PT   Novel regulator of G-protein signaling nucleic acids and polypeptides,
PT   useful as diagnostic and investigative tools and to treat G-protein
PT   signaling disorders.
XX
PS   Claim 8; Page 102; 105pp; English.
XX
CC   The RGS (regulators of G-protein signaling) protein genes, clones
CC   AAH16395 and m1975, were identified in human and murine spleen cDNA
CC   libraries, respectively. Both proteins have unique N- and C-terminal
CC   sequences. The C-terminal location of the RGS domain is consistent with
CC   RGSs known to act as GTPase activating proteins (GAPs) for G-alpha
CC   proteins. G-alpha-i linked receptors support rapid adhesion and directed
CC   migration of leukocytes and other cell types. The novel RGS proteins may
CC   be used to modulate cell adhesion and chemotaxis, e.g. for aiding wound
CC   repair. The RGS proteins, related cDNAs and anti-RGS antibodies are
CC   useful for modulation, diagnosis and treatment of immune and respiratory
CC   disorders
XX
SQ   Sequence 235 AA;

```
Query Match          85.0%;    Score 1032;  DB 3;  Length 235;
Best Local Similarity  83.8%;  Pred. No. 2.7e-90;
Matches   197;  Conservative  19;  Mismatches  19;  Indels   0;  Gaps   0;

Qy    1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60
          |:.:|:|::||||||||||||||:||:||||||||::|||||||||||||:.|:|::
Db    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
          |||:||..|||||||||||||||:|||||||||||||:.||||||||||||||||:|||
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy  121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
          ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
```

Matches 197; Conservative 19; Mismatches 19; Indels 19; Gaps 0;

Qy   1 MDMSLVFFSQLNMCESEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60
        |:||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 METTLLFFSQINMCESEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPBFHED 60

Qy   61 TQASRSALLAKETRVSPEEAVKWAEESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
        |::||| |||||||||||||||||| |||||||||:||||||||||||||||||||||:|
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSENIEFWIACE 120

Qy 121 DFKKCKBPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
        ||||| ||||||||||||||||||||:|||||||||||||||:||||||||||:||||||
Db 121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy 181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
        |||||||||||||||||| |:|||||||||||||||||||| |:|:|||||||||
Db 181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235

RESULT 5
AAU10749
ID   AAU10749 standard; protein; 235 AA.
XX
AC   AAU10749;
XX
DT   12-MAR-2002 (first entry)
XX
DE   Human RGS18 polypeptide.
XX
KW   Human; regulator of G protein signalling; RGS18; arterial thrombosis;
KW   platelet activation dysfunction; myocardial infarction; stroke;
KW   coronary artery disease; cerebrovascular disease; unstable angina;
KW   deep vein thrombosis; systemic thromboembolism; anti-coagulant;
KW   invasive cardiac procedure.
XX
OS   Homo sapiens.
XX
PN   WO200183514-A2.
XX
PD   08-NOV-2001.
XX
PF   26-APR-2001; 2001WO-US013540.
XX
PR   28-APR-2000; 2000US-0200786P.
PR   02-AUG-2000; 2000GB-00018833.
XX
PA   (AVET) AVENTIS PHARM PROD INC.
XX
PI   Murray DL, Gagnon AW;
XX
DR   WPI; 2002-055453/07.
DR   N-PSDB; AAS18340.
XX
PT   Isolated regulator of G protein signaling polypeptide, useful for
PT   prevention/treatment of platelet activation dysfunction, such as arterial
PT   thrombosis, myocardial infarction, coronary artery disease and stroke.
XX
PS   Claim 31; Fig 1; 127pp; English.
XX
CC   The present invention relates to the isolation of polynucleotide
CC   sequences, that encode a novel regulator of G protein signalling (RGS)
CC   polypeptide, RGS18, from human platelets. The invention also provides
CC   nucleotide primers and probes specific for an RGS18 nucleic acid. The
CC   sequences of the invention are useful for the manufacture of a medicament
CC   for the prevention or treatment of a platelet activation dysfunction,
CC   such as arterial thrombosis, myocardial infarction, coronary artery
CC   disease, cerebrovascular disease, unstable angina, deep vein
CC   thrombosis, systemic thromboembolism, and in invasive cardiac procedures
CC   for anti-coagulant purposes. The sequences of the invention can also be
CC   used for the detection of RGS18 nucleic acids and RGS18 polypeptides. The
CC   present sequence represents the RGS18 polypeptide isolated from human
CC   platelets
XX

Db 121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy 181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
        |||||||||||||||||| |:|||||||||||||||||||| |:|:|||||||||
Db 181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235

RESULT 4
AAE25829
ID   AAE25829 standard; protein; 235 AA.
XX
AC   AAE25829;
XX
DT   15-NOV-2002 (first entry)
XX
DE   Human regulator of G-protein signalling (RGS) protein, h16395.
XX
KW   Human; screening; RGS; regulator of G-protein signalling; pancreatitis;
KW   inflammatory disorder; allergy; Grave's disease; arthritis; sinusitis;
KW   respiratory disorder; asthma; pneumonia; therapy; immune disorder;
KW   haematological disorder; haematopoiesis; platelet-associated disorder;
KW   thrombocytopaenia; migration; invasive disorder; leukaemia; anaemia;
KW   erythrocyte-associated disorder; hepatitis; h16395.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   Domain   82..201
FT   /note= "RGS domain"
XX
PN   US6410240-B1.
XX
PD   25-JUN-2002.
XX
DR   WPI; 2002-588886/63.
DR   N-PSDB; AAD42497.
XX
PF   04-FEB-2000; 2000US-00498959.
XX
PR   04-FEB-1999; 99US-00244314.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Hodge MR, Yowe D;
XX
PS   Claim 1; Col 55-56; 42pp; English.
XX
CC   The invention relates to screening assays for identifying an agent that
CC   binds to: human or mouse RGS (regulators of G-protein signalling) protein
CC   or its variant, or RGS protein encoded by homologous DNA sequences; or a
CC   host cell expressing the RGS protein or its variant, or RGS protein
CC   encoded by homologous DNA sequences. The invention is useful for
CC   identifying an agent that binds to human or mouse RGS protein or its
CC   variant, or RGS protein encoded by homologous DNA sequences, or a host
CC   cell expressing the RGS protein or its variant, or RGS protein encoded by
CC   homologous DNA sequences. The agents identified using the invention are
CC   useful for modulating the activity of RGS proteins and thus useful for
CC   treating immune and inflammatory disorders (Grave's disease, allergy,
CC   arthritis), respiratory disorders (asthma, pneumonia, sinusitis),
CC   haematological disorders (haematopoiesis, migration), platelet-associated
CC   disorders (thrombocytopaenia), invasive disorders (leukaemia),
CC   erythrocyte-associated disorders (anaemia), pancreatitis, hepatitis etc.
CC   The present sequence is human RGS protein, h16395
XX
SQ   Sequence 235 AA;

Query Match   85.0%; Score 1032; DB 5; Length 235;
Best Local Similarity   83.8%; Pred. No. 2.7e-90;

```
SQ  Sequence 235 AA;

Query Match          85.0%;     Score 1032;  DB 5;   Length 235;
Best Local Similarity  83.8%;   Pred. No. 2.7e-90;
Matches 197;  Conservative  19;  Mismatches  19;  Indels  0;  Gaps  0;

Qy   1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60
        |:::||||||:|||||||||||||:|||||||||:|:||||||||||||||:|::|| :
Db   1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy  61 TQASRSALLAKETRVSPEEAVKWAAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
        |::||| ||||||||||||||||| :|||||||||||| ::||||||||||||||||:|||
Db  61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy 121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
        ||||.| .||| :|||||||||||| ||||||| |||||||:  | ||||||||:|||||
Db 121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy 181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
        |||||:||| ||||| | | :|:||||||||||||||||||  :|||:|||||||
Db 181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235


RESULT 6
ABP64801
ID  ABP64801 standard; protein; 235 AA.
XX
AC  ABP64801;
XX
DT  25-FEB-2003  (first entry)
XX
DE  Human protein SEQ ID 461.
XX
KW  Human; expressed sequence tag; EST; haematopoietic disorder;
KW  central nervous system disease; viral infection;
KW  peripheral nervous system disease; non-healing wound; infectious disease;
KW  immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW  fungal infection; autoimmune disorder; coagulation disorder; nootropic;
KW  antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW  cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW  immunostimulant; cerebroprotective.
XX
OS  Homo sapiens.
XX
PN  WO200259260-A2.
XX
PD  01-AUG-2002.
XX
PF  16-NOV-2001; 2001WO-US042950.
XX
PR  17-NOV-2000; 2000US-00714936.
XX
PA  (HYSE-) HYSEQ INC.
XX
PI  Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA,
PI  Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR  WPI; 2002-590824/63.
DR  N-PSDB; ABQ99387.
XX
PT  New isolated polynucleotide, useful in research, diagnostic or
PT  therapeutic methods, e.g. preventing or treating disorders involving
PT  aberrant protein expression or biological activity.
XX
PS  Claim 20; SEQ ID NO 461; 394pp; English.
XX
CC  The present invention relates to novel human coding sequences (ABQ99268-
CC  ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC  therapeutic, diagnostic and research methods. The polynucleotides may be
CC  used in the field of molecular biology as hybridisation probes, primers
CC  for PCR, for chromosome and gene mapping, for the recombinant production
CC  of protein, or in generation of anti-sense DNA or RNA. The
CC  polynucleotides are useful in diagnostics as expressed sequence tags
CC  (ESTs) for identifying expressed genes or for physical mapping of the
CC  human genome. The proteins may be used as molecular weight markers, or as
CC  nutritional sources or supplements. The proteins may be used to maintain
CC  and expand cell population in a totipotential or pluripotential state
CC  useful for re-engineering damaged or diseased tissues, transplantation,
CC  manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC  polynucleotides and proteins are useful for preventing, treating or
CC  ameliorating disorders involving aberrant protein expression or
CC  biological activity, e.g. haematopoietic disorders, central/peripheral
CC  nervous system diseases, mechanical and traumatic disorders, non-healing
CC  wounds, immune deficiencies and disorders, infectious diseases caused by
CC  viral, bacterial or fungal infection, autoimmune disorders, allergic
CC  reactions and conditions, coagulation disorders, or cancer. The
CC  polynucleotide sequences of the invention were assembled from ESTs
CC  isolated mainly by sequencing by hybridisation, and in some cases,
CC  sequences obtained from one or more public databases. Note: The sequence
CC  data for this patent did not form part of the printed specification, but
CC  was obtained in electronic format directly from WIPO at
CC  ftp.wipo.int/pub/published_pct_sequences
XX
SQ  Sequence 235 AA;

Query Match          85.0%;     Score 1032;  DB 5;   Length 235;
Best Local Similarity  83.8%;   Pred. No. 2.7e-90;
Matches 197;  Conservative  19;  Mismatches  19;  Indels  0;  Gaps  0;

Qy   1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60
        |:::||||||:|||||||||||||:|||||||||:|:||||||||||||||:|::|| :
Db   1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy  61 TQASRSALLAKETRVSPEEAVKWAAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
        |::||| ||||||||||||||||| :|||||||||||| ::||||||||||||||||:|||
Db  61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy 121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
        ||||.| .||| :|||||||||||| ||||||| |||||||:  | ||||||||:|||||
Db 121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy 181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
        |||||:||| ||||| | | :|:||||||||||||||||||  :|||:|||||||
Db 181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235


RESULT 7
ABP64054
ID  ABP64054 standard; protein; 227 AA.
XX
AC  ABP64054;
XX
DT  04-NOV-2002  (first entry)
XX
DE  Human ORF424.
XX
KW  Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW  Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW  human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW  cancer; cardiovascular disease; allergy; autoimmune disease;
KW  wound healing; blood coagulation disorder; inflammatory disorder.
XX
OS  Homo sapiens.
XX
PN  US2002082206-A1.
XX
PD  27-JUN-2002.
XX
PF  30-MAY-2001; 2001US-00867550.
PR  30-MAY-2000; 2000US-0208427P.
XX
PA  (LEAC/) LEACH M D.
PA  (MEHR/) MEHRABAN F.
PA  (CONL/) CONLEY P B.
PA  (TOPP/) TOPPER J N.
```

PA (1AWD/) LAW D.
XX
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
DR WPI; 2002-626554/67.
DR N-PSDB; ABQ99617.
XX
PT New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.
XX
PS Claim 10; SEQ ID NO 848; 78pp; English.
XX
CC The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/sequence.html?DocID=20020082205
XX
SQ Sequence 227 AA;

Query Match   81.9%; Score 994; DB 5; Length 227;
Best Local Similarity   83.7%; Pred. No. 1.1e-86;
Matches 190; Conservative 18; Mismatches 19; Indels   0; Gaps   0;

Qy   1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60
         : ||:|||||||||||||||||||||||||||:|||||||||||||||||||:| :|| :
Db   1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLIVQKPEFHED 60

Qy   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
         |||:|||||||||||||||||||||||| :|||||||| |||||||||||||||||||:|||
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy 121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
         |:||| ||||||||||||||| :||||||||||||||::|||||||||||||||||||||
Db 121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVIINSITQPTLHSFDAAQSRV 180

Qy 181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDV 227
         |||| ||:||||||| :|||:||||||||||||||||||| :||||
Db 181 YQLMEQDSVTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDV 227

RESULT 8
ABP96779
ID ABP96779 standard; protein; 211 AA.
XX
AC ABP96779;
XX
DT 05-JUN-2003 (first entry)
XX
DE Human COPD related protein SEQ ID NO.29.
XX
KW Human; chronic obstructive pulmonary disease; COPD; chronic lung disease.
XX
OS Homo sapiens.
XX
PN WO200297127-A2.
XX
PD 05-DEC-2002.
XX
PF 28-MAY-2002; 2002WO-EP005835.
XX
PR 31-MAY-2001; 2001GB-00013266.
XX

PA (FARB) BAYER AG.
XX
PI Oellers N, Gehrmann M, Kallabis H, Hall R, Schulze T, Kroegel C;
XX
DR WPI; 2003-140492/13.
DR N-PSDB; ACC46750.
XX
PT Predicting, diagnosing or prognosing chronic lung disease, by detecting a
PT chronic obstructive pulmonary disease (COPD) gene in a biological sample.
XX
PS Claim 8; Page 119-120; 214pp; English.
XX
CC The present invention describes a method for predicting, diagnosing or
CC prognosing chronic lung disease by detecting a chronic obstructive
CC pulmonary disease (COPD) gene related polynucleotide (see ACC46750 to
CC ACC46777) which encode the COPD related proteins in ABP96779 to
CC ABP98806). The method is useful for predicting, diagnosing or prognosing
CC chronic lung disease in a biological sample. The COPD genes and proteins
CC encoded by them from the present invention (I) can be used for treating
CC or preventing chronic lung disease in a mammal. (I) can be used in an
CC animal model for determining the efficacy, toxicity, or side effects of
CC treatment with (I), and determining the mechanism of action of (I).
CC ACC46778 to ACC46903 represent COPD related PCR primers and probes used
CC in an example from the present invention
XX
SQ Sequence 211 AA;

Query Match   35.4%; Score 429.5; DB 6; Length 211;
Best Local Similarity 44.3%; Pred. No. 9.7e-33;
Matches 93; Conservative 38; Mismatches 70; Indels   9; Gaps   3;

Qy   4 SLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIR--AKEKRNRLSLLLQRPDFHGE 60
         :|:|| :|| :||::||||||| :|: ||| :: | |:| ||||| || :||||
Db   3 SAMFLAVQHDCRPMDKS-----AGSGHKSEKREKMKRTLLKDWKTRLSYFLQNSSTPGK 57

Qy   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
         : ::|| | |::| |||| || |||||||||||| | || | : : |::|| ||
Db   58 PKTGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACE 117

Qy 121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
         |||| ||||| ||:|||||:|||:|||||||||||||| || |:| || || |:|
Db 118 DFKKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRV 177

Qy 181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPT 210
         |||||:||||||||| :|||||||:|: | :
Db 178 YSLMENNSYPRFLESEFYQDLCK-KPQITT 206

RESULT 9
ABP98020
ID ABP98020 standard; protein; 211 AA.
XX
AC ABP98020;
XX
DT 11-AUG-2003 (first entry)
XX
DE Amino acid sequence of regulator polypeptide of G-protein signalling 2.
XX
KW Human; regulator polypeptide of G-protein signalling; RGS; chromosome 1;
KW RGS2; asthma; diabetes.
XX
OS Homo sapiens.
XX
PN WO2003016345-A1.
XX
PD 27-FEB-2003.
XX
PF 16-AUG-2002; 2002WO-EP009173.
XX
PR 16-AUG-2001; 2001US-0312345P.
XX
PA (FARB) BAYER AG.
XX

```
PI   Koehler RH;
XX
DR   WPI; 2003-278547/27.
XX
PT   New polynucleotide encoding a regulator polypeptide of G-protein
PT   signaling (RGS) and RGS protein, useful in gene therapy to modulate the
PT   activity of RGS in a disease, particularly in asthma or diabetes.
XX
PS   Disclosure; Fig 3; 99pp; English.
XX
CC   The present sequence is a human regulator polypeptide of G-protein
CC   signaling (RGS) 2. The RGS gene is located on chromosome 1. The RGS
CC   polynucleotide and polypeptide are useful for identifying test compounds,
CC   that may act as agonists or antagonists at the receptor site and which
CC   can be regulated to provide therapeutic effects. The polynucleotide,
CC   polypeptide or vectors comprising the polynucleotide, and the reagents
CC   are useful for modulating the activity of RGS in a disease, particularly
CC   in asthma or diabetes
XX
SQ   Sequence 211 AA;

Query Match             35.4%;  Score 429.5; DB 6;  Length 211;
Best Local Similarity   44.3%;  Pred. No. 9.7e-33;
Matches   93;  Conservative   38;  Mismatches   70;  Indels    9;  Gaps    3;

Qy    4  SLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIR--AKEKRNRLSLLLQRPDFHGE   60
            :|::||  :|  :|: .:|||  :|| |::|:||
Db    3  SAMFLAVQHDCRPMDKS-----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSSTPGK  57

Qy   61  TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRPLKTEFSEENIEFWVACE  120
            ||.: | | :||||||  :||| :: ||||:|:  |::|  || ||: |:|||||||||||
Db   58  PKTGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACE  117

Qy  121  DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSPDTAQSRV  180
            ||||:|  ||.:|| | | | | :::|| ||:| ||||
Db  118  DFKKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRV  177

Qy  181  YQLMEHDSYKRFLKSETYLHLIEGRPQRPT  210
            |||:|| || || :|| :||
Db  178  YSLMENNSYPRFLESEFYQDLCK-KPQITT  206

RESULT 10
ADD14019
ID   ADD14019 standard; protein; 211 AA.
XX
AC   ADD14019;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Human src biomarker polypeptide SEQ ID NO:208.
XX
KW   predictor set; protein tyrosine kinase activity modulator;
KW   protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW   gene therapy; drug sensitivity; genetic profile; cancer; human.
XX
OS   Homo sapiens.
XX
PN   WO200306239S-A2.
XX
PD   31-JUL-2003.
XX
PF   17-JAN-2003; 2003WO-US001981.
XX
PR   18-JAN-2002; 2002US-0350061P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Huang F, Fairchild CR, Lee FY, Shaw P;
XX
DR   WPI; 2003-636735/60.
DR   N-PSDB; ADD14613.
XX
```

```
PT   New polynucleotides and polypeptides for predicting the activity of
PT   compounds that interact with protein tyrosine kinases and/or protein
PT   tyrosine kinase pathways.
XX
PS   Claim 10; SEQ ID NO 208; 139pp; English.
XX
CC   The present invention describes a predictor set comprising a plurality of
CC   polynucleotides or polypeptides whose expression pattern is predictive of
CC   the response of cells to treatment with a compound that modulates protein
CC   tyrosine kinase activity or members of the protein tyrosine kinase
CC   pathway. Also described: (1) predicting whether a compound is capable of
CC   modulating the activity of cells, comprising obtaining a sample of cells,
CC   determining whether the cells express a plurality of markers, and
CC   correlating the expression of the markers to the compound's ability to
CC   modulate the activity of the cells; (2) a plurality of cell lines for
CC   identifying polynucleotides and polypeptides whose expression levels
CC   correlate with compound sensitivity or resistance of cells associated
CC   with a disease state; and (3) identifying polynucleotides and
CC   polypeptides that predict compound sensitivity or resistance of cells
CC   associated with a disease state, comprising subjecting the plurality of
CC   cell lines to one or more compounds, analysing the expression pattern of
CC   a microarray of polynucleotides or polypeptides, and selecting
CC   polynucleotides or polypeptides that predict the sensitivity or
CC   resistance of cells associated with a disease state by using the
CC   expression pattern of the microarray. The polynucleotides and
CC   polypeptides have cytostatic activities, and can be used in gene therapy.
CC   The polynucleotides and polypeptides are useful in predicting the
CC   activity of compounds that interact with protein tyrosine kinases and/or
CC   protein tyrosine kinase pathway. These may be used in determining drug
CC   sensitivity in patients to allow the development of individualized
CC   genetic profiles which aid in treating diseases and disorders (e.g.
CC   cancer) based on patient response at a molecular level. The present
CC   sequence is used in the exemplification of the present invention.
XX
SQ   Sequence 211 AA;

Query Match             35.4%;  Score 429.5; DB 7;  Length 211;
Best Local Similarity   44.3%;  Pred. No. 9.7e-33;
Matches   93;  Conservative   38;  Mismatches   70;  Indels    9;  Gaps    3;

Qy    4  SLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIR--AKEKRNRLSLLLQRPDFHGE   60
            :|::||  :|  :|: .:|||  :|| |::|:||
Db    3  SAMFLAVQHDCRPMDKS-----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSSTPGK  57

Qy   61  TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRPLKTEFSEENIEFWVACE  120
            ||.: | | :||||||  :||| :: ||||:|:  |::|  || ||: |:|||||||||||
Db   58  PKTGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACE  117

Qy  121  DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSPDTAQSRV  180
            ||||:|  ||.:|| | | | | :::|| ||:| ||||
Db  118  DFKKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRV  177

Qy  181  YQLMEHDSYKRFLKSETYLHLIEGRPQRPT  210
            |||:|| || || :|| :||
Db  178  YSLMENNSYPRFLESEFYQDLCK-KPQITT  206

RESULT 11
AAB56929
ID   AAB56929 standard; protein; 220 AA.
XX
AC   AAB56929;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Human prostate cancer antigen protein sequence SEQ ID NO:1507.
XX
KW   Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW   neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW   vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW   antibacterial; gene therapy; neural; immune; reproductive; renal;
KW   gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW   wound; infectious disease.
```

XX
OS   Homo sapiens.
XX
PN   WO20005174-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US005988.
XX
PR   12-MAR-1999; 99US-0124270P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (ROSE/) ROSEN C A.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2000-587513/55.
DR   N-PSDB; AAF16132.
XX
PT   Prostate cancer associated gene sequences, referred to as prostate cancer
PT   antigens, useful for treatment, prevention, and diagnosis of disorders
PT   such as prostate cancer.
XX
PS   Claim 11; Page 1946-1947; 2338pp; English.
XX
CC   AAF15566 to AAF16505 encode the human prostate cancer associated
CC   proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC   The prostate cancer antigens can have neuroprotective, cytostatic,
CC   cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC   nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC   and can be used in gene therapy. The prostate cancer antigen
CC   polynucleotides may be used for detection of prostate cancer, chromosome
CC   identification, as chromosome markers, and for numerous other diagnostic
CC   or research purposes. The prostate cancer antigens may be used to treat
CC   disorders such as neural, immune, muscular, reproductive,
CC   gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC   disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC   AAB57303 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 220 AA;

Query Match   35.4%; Score 429.5; DB 3; Length 220;
Best Local Similarity   44.3%; Pred. No. 1e-32;
Matches   93; Conservative   38; Mismatches   70; Indels   9; Gaps   3;

Qy   4 SLVFFSQLMCESKEKTFFKLMHGSGKETSIEAKIR---AKEKRNLSLLQRPDFHGE 60
         | ||::|||:::::::|||||| |:|: |||| | : ||::|| | : |||
Db   12 SAMFLAVQHDCRPMDKS----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSSTPGK 66

Qy   61 TQASRSALLAKETRVSPEEAVKWAESPDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
         | ||:| | |::|| :||| | |::||| ::|::|::|: |||| ||:||||||||:||
Db   67 PKTGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACE 126

Qy 121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
         ||| | |||::| |||||||| ||:|||:|||||:||||||||||||||||||:||:||
Db 127 DFKKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQKTLIAQNIQEATSGCFTTAQKRV 186

Qy 181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPT 210
         || |::|:||||:|:::| |||| : ||
Db 187 YSLMENNSYPRFLESEFYQDLCK-KPQITT 215

RESULT 12
AAW62075
ID   AAW62075 standard; protein; 181 AA.
XX
AC   AAW62075;
XX
DT   14-SEP-1998 (first entry)
XX
DE   Human regulator of G-protein signalling.
XX

KW   Human; regulator; G-protein signalling; HRGS; cancer; inflammation;
KW   hypertension; cardiovascular shock; arrythmias; asthma.
XX
OS   Homo sapiens.
XX
PN   WO982012B-A1.
XX
PD   14-MAY-1998.
XX
PF   06-NOV-1997; 97WO-US018476.
XX
PR   08-NOV-1996; 96US-00748483.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Hillman JL, Goli SK;
XX
DR   WPI; 1998-286944/25.
DR   N-PSDB; AAV38084.
XX
PT   Regulator of G-protein signalling - used to develop products for treating
PT   e.g. cancer, inflammation, hypertension, cardiovascular shock,
PT   arrhythmias or asthma.
XX
PS   Claim 1; Fig 1; 66pp; English.
XX
CC   The present sequence represents human regulator of G-protein signalling
CC   (HRGS). The HRGS regulates G-protein signalling in cancer cells and may
CC   be useful in the treatment of any cancer, especially cancers of the brain
CC   and thyroid. Products of the present invention can also be used for
CC   treating other conditions associated with uncontroled cell signalling
CC   such as inflammation. The products can also be used to modulate HRGS
CC   activity in response to disorders involving the sympathetic nervous
CC   system including hypertension, cardiovascular shock, arrhythmias and
CC   asthma. The products can also be used for detection, diagnosis and drug
CC   screening
XX
SQ   Sequence 181 AA;

Query Match   35.2%; Score 427.5; DB 2; Length 181;
Best Local Similarity   50.3%; Pred. No. 1.2e-32;
Matches   84; Conservative   33; Mismatches   47; Indels   3; Gaps   1;

Qy   40 RAKEKRNLSLLLQRPDFHGE---TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDG 96
         ||||:||:|||||:||||| :|||||| ||:||| :||| | |::|||||||||
Db   15 RAKEIKIKLGILLQKPDSVGDLVIPVNEKPEKPAXTQKTSLDEALQWRDSLDKLLQNNYG 74

Qy   97 VDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFH 156
         ||||| ||||:|||||||||:|||:|| | |||::|| |||||||| ||:|||:|||||
Db   75 LASFKSFLKSEFSEENLEFWIACEDYKKIKSPAKMAEKAKQIYEEFIQTEAPKEVNIDHF 134

Qy 157 TKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIE 203
         |||||||||||||||||:||:||| || |::|:||||:|:::| ||
Db 135 TKDITMNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK 181

RESULT 13
AAY53931
ID   AAY53931 standard; protein; 181 AA.
XX
AC   AAY53931;
XX
DT   13-MAR-2000 (first entry)
XX
DE   A human regulator of G protein signalling RGS5 protein.
XX
KW   Activator of G protein signalling; AGS; ras-related G protein;
KW   GTP hydrolysis; G protein activity; pheromone response pathway;
KW   G protein-coupled signal transduction; G-gamma selectivity;
KW   cellular signal transduction; regulator of G protein signalling; RGS5.
XX
OS   Homo sapiens.
XX

```
PN   WO9958670-A1.
XX
PD   18-NOV-1999.
XX
PF   07-MAY-1999;    99WO-US010151.
XX
PR   08-MAY-1998;    98US-0084842P.
PR   07-OCT-1998;    98US-0103355P.
XX
PA   (CADU-) CADUS PHARM CORP.
XX
PI   Cismowski M,  Duzic E;
XX
DR   WPI; 2000-072337/06.
DR   N-PSDB; AA236910.
XX
PT   A new activator of G protein signalling used to treat disorders
PT   characterized by an aberrant AGS protein activity.
XX
PS   Claim 74; Page 141-141; 162pp; English.
XX
CC   The present sequence represents a regulator of G protein signalling (AGS)
CC   protein, RGS5. The specification also describes an activator of G protein
CC   signalling (AGS) protein. The AGS cDNA sequence was isolated from a human
CC   liver cDNA library. The AGS protein exhibits homology to ras-related G
CC   proteins, and contains alterations in conserved amino acids consistent
CC   with a deficiency in GTP hydrolysis activity. AGS stimulates G protein
CC   activity, G protein-coupled signal transduction and the pheromone
CC   response pathway in a receptor-independent manner. The AGS protein also
CC   shows G-gamma selectivity, as measured by growth assays in yeast
CC   expressing various mammalian G-gamma constructs, and tissue-specific
CC   expression, as measured by Northern blot analysis. The AGS protein can be
CC   used to screen for compounds that modulate cellular signal transduction.
CC   The protein is used to treat disorders characterized by an aberrant AGS
CC   protein activity or AGS nucleic acid expression
XX
SQ   Sequence 181 AA;

Query Match             35.2%;   Score 427.5;  DB 3;  Length 181;
Best Local Similarity   50.3%;   Pred. No. 1.2e-32;
Matches    84;  Conservative   33;  Mismatches   47;  Indels    3;  Gaps     1;

Qy         40 RAKEKRNRLSLLLQRPDFHGE--TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDG 96
              ||||.||||:.:|||:||| |    :||||:||:::||||::| ||||| :
Db         15 RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG 74

Qy         97 VDAFTRFLKTEFSBENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFH 156
              ::||:||||.:|||:|||:|||:::|||| ||::|:::::|| |||::||||||||||
Db         75 LASFKSFLKSEFSEENLEFWIACEDYKKIKSPAKMAEKAKQIYEEFIQTEAPKEVNIDHF 134

Qy        157 TKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIE 203
              ||| ||:| |||||:||:|:| |:|||:|:||||| ::| |::||
Db        135 TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK 181

RESULT 14
ABP62144
ID   ABP62144 standard; protein; 190 AA.
XX
AC   ABP62144;
XX
DT   12-NOV-2002 (first entry)
XX
DE   Human secreted protein SEQ ID NO 197.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
```

```
OS   Homo sapiens.
XX
PN   WO20257420-A2.
XX
PD   25-JUL-2002.
XX
PF   17-JAN-2002; 2002WO-US001109.
XX
PR   18-JAN-2001; 2001US-0262066P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Moore PA,  Ruben SM,  Lafleur DW,  Shi Y,  Rosen CA,  Olsen H;
PI   Ebner R,  Brewer LA;
XX
DR   WPI; 2002-599716/64.
XX
PT   New polynucleotides and polypeptides useful for diagnosing, prognosing,
PT   treating or preventing e.g. neurodegenerative, central nervous system,
PT   autoimmune, respiratory, reproductive, or inflammatory diseases or
PT   disorders.
XX
PS   Claim 11; Page 71; 785pp; English.
XX
CC   The invention relates to novel genes (ABQ2553-ABQ2607) and proteins
CC   (ABP62013-ABP62153) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC   breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
CC   infections
XX
SQ   Sequence 190 AA;

Query Match             35.2%;   Score 427.5;   DB 5;   Length 190;
Best Local Similarity   50.3%;   Pred. No. 1.3e-32;
Matches    84;  Conservative   33;  Mismatches   47;  Indels    3;  Gaps     1;

Qy         40 RAKEKRNRLSLLLQRPDFHGE--TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDG 96
              ||||.||||:.:|||:||| |    :||||:||:::||||::| ||||| :
Db         24 RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG 83

Qy         97 VDAFTRFLKTEFSBENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFH 156
              ::||:||||.:|||:|||:|||:::|||| ||::|:::::|| |||::||||||||||
Db         84 LASFKSFLKSEFSEENLEFWIACEDYKKIKSPAKMAEKAKQIYEEFIQTEAPKEVNIDHF 143

Qy        157 TKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIE 203
              ||| ||:| |||||:||:|:| |:|||:|:||||| ::| |::||
Db        144 TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK 190

RESULT 15
ABB83788
ID   ABB83788 standard; protein; 181 AA.
XX
AC   ABB83788;
XX
DT   14-AUG-2002 (first entry)
XX
DE   Rat disease associated protein SEQ ID NO 1.
XX
KW   Rat; cardiant; heart disease; cardiovascular disease; cardiac infarction;
KW   argina; gene therapy.
XX
OS   Rattus sp.
XX
```

```
XX
PN   WO200233082-A1.
XX
PD   25-APR-2002.
XX
PF   18-OCT-2001; 2001WO-JP009140.
XX
PR   19-OCT-2000; 2000JP-00319912.
PR   16-NOV-2000; 2000JP-00350183.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Koyama N, Tanida S, Watanabe T;
XX
DR   WPI; 2002-394557/42.
DR   N-PSDB; ABL99975.
XX
PT   Disease-associated gene and encoded RGS5-like protein, applicable in
PT   diagnosis and prevention or treatment of heart diseases e.g.
PT   cardiovascular diseases, cardiac infarction, heart failure and angina,
PT   including gene therapy.
XX
PS   Claim 1; Page 71-72; 82pp; Japanese.
XX
CC   The invention relates to a protein (ABB83788) or its salt. The protein
CC   and encoded DNA (ABL99975) are applicable in diagnosis and prevention or
CC   treatment of heart diseases e.g. cardiovascular diseases, cardiac
CC   infarction, heart failure and angina, including gene therapy
XX
SQ   Sequence 181 AA;

Query Match             34.0%;  Score 412.5;  DB 5;  Length 181;
Best Local Similarity   49.1%;  Pred. No. 3.3e-31;
Matches   86;  Conservative  27;  Mismatches  41;  Indels  21;  Gaps   2;

Qy     40 RAKEKRNRLSLLLQRPDFHGETQASRSAL----------LAKETRVSPEEAVKWAESF 87
          |||||:|:|||||:||                ||::||::||:|:|
Db     15 RAKEIKILGILLQKPD--------SAVDLVIPYNEKPEKPAKAHKPSLEEVLQWRQSL 65

Qy     88 DKLLSHRDGVDAFTRPLKTEFSEENIEFWVACEDFKKCKEPCQIILKAKAIYEKFIQNDA 147
          ||||||||| ||||||:||:|||:|||||||||||||:|| :||| |||:||:|
Db     66 DKLLQSNYGFASFKSFLKSEFSEENLEFWVACENYKIKSPIKMAEKAKQIYEEFIQTEA 125

Qy    148 PKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLI 202
          |||||||| ||::|:::|:|||||||||||||:|||||:||||||:||:||:|||
Db    126 PKEVNIDFTKDITMKNLVEPSPHSFDLAQKRIYALMEKDSLPRFVRSEFYKELI 180
```

Search completed: August 20, 2004, 16:44:15
Job time : 80 secs

```
                    GenCore version 5.1.6
         Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    August 20, 2004, 16:40:58 ; Search time 18.5 Seconds
                                        (without alignments)
                                        655.789 Million cell updates/sec

Title:         US-09-894-749-4
Perfect score: 1214
Sequence:      1 MDMSLVFFSQLNMCESKEKT..........SRSFTYNDFQDVKSDVAIWL 235

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:      389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters:     389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued Patents AA:*
              1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
              2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
              3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
              4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
              5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
              6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

              Pred. No. is the number of results predicted by chance to have a
              score greater than or equal to the score of the result being printed,
              and is derived by analysis of the total score distribution.

                                 SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1214	100.0	235	3	US-09-244-314-4	Sequence 4, Appli
2	1214	100.0	235	4	US-09-498-959-4	Sequence 4, Appli
3	1032	85.0	235	3	US-09-244-314-2	Sequence 2, Appli
4	1032	85.0	235	4	US-09-498-959-2	Sequence 2, Appli
5	429.5	35.4	211	2	US-08-748-483-4	Sequence 4, Appli
6	427.5	35.2	181	2	US-08-748-483-1	Sequence 1, Appli
7	397.5	32.7	205	2	US-08-829-110-6	Sequence 6, Appli
8	397.5	32.7	205	2	US-08-748-483-5	Sequence 5, Appli
9	370	30.5	120	1	US-08-890-865A-11	Sequence 11, Appli
10	364	30.0	202	2	US-08-274-318-2	Sequence 2, Appli
11	364	30.0	202	2	US-08-754-108-2	Sequence 2, Appli
12	363	29.9	202	2	US-08-463-081B-2	Sequence 2, Appli
13	363	29.9	202	2	US-08-461-379A-2	Sequence 2, Appli
14	363	29.9	202	2	US-08-462-390B-2	Sequence 2, Appli
15	363	29.9	202	3	US-08-870-815-4	Sequence 4, Appli
16	363	29.9	202	3	US-08-463-074B-2	Sequence 2, Appli
17	363	29.9	202	3	US-08-465-585C-2	Sequence 2, Appli
18	363	29.9	202	2	US-08-652-446-2	Sequence 2, Appli
19	363	29.9	202	3	US-08-949-004-4	Sequence 4, Appli
20	359.5	29.6	121	2	US-08-588-258B-32	Sequence 32, Appli
21	359.5	29.6	121	2	US-08-460-505-32	Sequence 32, Appli
22	359.5	29.6	121	5	PCT-US96-08295-32	Sequence 32, Appli
23	354	29.2	120	4	US-08-890-865A-12	Sequence 12, Appli
24	348	28.7	120	4	US-08-890-865A-14	Sequence 14, Appli
25	348	28.7	201	2	US-08-726-228-2	Sequence 2, Appli
26	348	28.7	201	3	US-08-870-815-2	Sequence 2, Appli
27	348	28.7	201	3	US-08-949-004-2	Sequence 2, Appli
28	347.5	28.6	196	2	US-08-829-110-5	Sequence 5, Appli
29	347.5	28.6	196	2	US-08-748-483-3	Sequence 3, Appli
30	347.5	28.6	196	4	US-09-702-705-339	Sequence 339, App
31	347.5	28.6	196	4	US-09-736-457-339	Sequence 339, App
32	347.5	28.6	196	4	US-09-614-124B-339	Sequence 339, App
33	347.5	28.6	196	4	US-09-671-32-339	Sequence 339, App
34	347.5	28.6	196	4	US-09-589-184-339	Sequence 339, App
35	336.5	27.7	119	2	US-08-588-258B-31	Sequence 31, Appl
36	336.5	27.7	119	2	US-08-460-505-31	Sequence 31, Appl
37	336.5	27.7	119	5	PCT-US96-08295-31	Sequence 31, Appl
38	334.5	27.6	119	4	US-08-890-865A-15	Sequence 15, Appl
39	328	27.0	120	4	US-08-890-865A-13	Sequence 13, Appl
40	307	25.3	217	4	US-09-976-594-51	Sequence 51, Appl
41	303	25.0	159	2	US-08-829-110-1	Sequence 1, Appli
42	284	23.4	120	4	US-08-890-865A-16	Sequence 16, Appl
43	277.5	22.9	243	2	US-08-829-110-3	Sequence 3, Appli
44	274.5	22.6	119	2	US-08-588-258B-30	Sequence 30, Appl
45	274.5	22.6	119	3	US-08-460-505-30	Sequence 30, Appl

```
                                 ALIGNMENTS

RESULT 1
US-09-244-314-4
; Sequence 4, Application US/09244314
; Patent No. 6274362
; GENERAL INFORMATION:
;  APPLICANT: Hodge, Martin R.
;  APPLICANT: Yowe, David
;  TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
;  FILE REFERENCE: 5800-19, 035800/174680
;  CURRENT APPLICATION NUMBER: US/09/244,314
;  CURRENT FILING DATE: 1999-02-04
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;  LENGTH: 235
;  TYPE: PRT
;  ORGANISM: Mus sp.
US-09-244-314-4

Query Match             100.0%;  Score 1214;  DB 3;  Length 235;
Best Local Similarity   100.0%;  Pred. No. 2.1e-115;
Matches  235;  Conservative   0;  Mismatches   0;  Indels    0;  Gaps     0;

Qy          1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLQRPDFHGE 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLQRPDFHGE 60

Qy         61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120

Qy        121 DFFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DFFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180

Qy        181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235

RESULT 2
US-09-498-959-4
; Sequence 4, Application US/09498959
; Patent No. 6410240
; GENERAL INFORMATION:
;  APPLICANT: Hodge, Martin R.
;  APPLICANT: Yowe, David
;  TITLE OF INVENTION: RGS-Containing Molecules and Uses
;  TITLE OF INVENTION: Thereof
;  FILE REFERENCE: 5800-19A
```

RESULT 4
US-09-498-959-2
; Sequence 2, Application US/09498959
; Patent No. 6410240
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; APPLICANT: Yowe, David
; TITLE OF INVENTION: RGS-Containing Molecules and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 5800-19A
; CURRENT APPLICATION NUMBER: US/09/498,959
; CURRENT FILING DATE: 2000-02-04
; EARLIER APPLICATION NUMBER: 09/244,314
; EARLIER FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-498-959-2

Query Match   85.0%; Score 1032; DB 4; Length 235;
Best Local Similarity   83.8%; Pred. No. 6.3e-97;
Matches   197; Conservative   19; Mismatches   19; Indels   0; Gaps   0;

```
Qy    1 MDMSLIVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60
          : :|:||||||:|||||||||||:||||||||||||||||||||||||||:|||:
Db    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFEHD 60

Qy   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
          :|:: |||||||||||||||||:||||||||||||:|||||||||||||||||||:
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy  121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
          ||||:|:|||||||||||||||||:|||:||||||:||||||:::|||||||||:|||
Db  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
          |||||:||:|||||::|| |:||||||||||||||||||| :||||:|||||||
Db  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
```

RESULT 5
US-08-748-483-4
; Sequence 4, Application US/08748483
; Patent No. 5955314
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: A NOVEL REGULATOR OF CELL SIGNALLING
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,483
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.

; CURRENT APPLICATION NUMBER: US/09/498,959
; CURRENT FILING DATE: 2000-02-04
; EARLIER APPLICATION NUMBER: 09/244,314
; EARLIER FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-498-959-4

Query Match   100.0%; Score 1214; DB 4; Length 235;
Best Local Similarity 100.0%; Pred. No. 2.1e-115;
Matches   235; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1 MDMSLIVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MDMSLIVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60

Qy   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120

Qy  121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180

Qy  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
```

RESULT 3
US-09-244-314-2
; Sequence 2, Application US/09244314
; Patent No. 6274362
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; APPLICANT: Yowe, David
; TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
; FILE REFERENCE: 5800-19, 035800/174680
; CURRENT APPLICATION NUMBER: US/09/244,314
; CURRENT FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-244-314-2

Query Match   85.0%; Score 1032; DB 3; Length 235;
Best Local Similarity   83.8%; Pred. No. 6.3e-97;
Matches   197; Conservative   19; Mismatches   19; Indels   0; Gaps   0;

```
Qy    1 MDMSLIVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60
          : :|:||||||:|||||||||||:||||||||||||||||||||||||||:|:  :
Db    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFEHD 60

Qy   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
          |:|: |||||||||||||||||:||||||||||||:|||||||||||||||||||:
Db   61 TRSSRGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy  121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
          ||||:|:|||||||||||||||||:|||:||||||||||||:::|||||||||:|||
Db  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNIDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
          |||||:||:|||||::|| |:||||||||||||||||||| :||||:|||||||
Db  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
```

```
;       REGISTRATION NUMBER: 36,749
;       REFERENCE/DOCKET NUMBER: PF-0157 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 415-855-0555
;       TELEFAX: 415-845-4166
;       TELEX:
; INFORMATION FOR SEQ ID NO: 4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 211 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     IMMEDIATE SOURCE:
;       LIBRARY: GenBank
;       CLONE: 292037
US-08-748-483-4

Query Match         35.4%;  Score 429.5; DB 2; Length 211;
Best Local Similarity 44.3%;  Pred. No. 8.3e-36;
Matches   93;  Conservative   38; Mismatches   70;  Indels    9;  Gaps    3;

Qy    4 SLVFFSQLNMCESEKTFFKLMHGSGKETSIEAKIR---AKEKRNRLSLLLQRPDFHGE 60
          |:| | ||: |  |::  :||| :: ||||||||||||
Db    3 SAMFLAVQHDCRPMDKS-----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSSTPGK 57

Qy   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
         ||||||  :|:|:|  ||  :||  : |  :|  |||||||
Db   58 PKTGKKSKQQAFIKPSPEEAQLWSEAFDEILASKYGLAAFRAFLKSEFCEENIEFWLACE 117

Qy  121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
         ||||  |||||:|| ||| :|  ||:|||:|
Db  118 DFKKTKSPQKLSSKARKIYTDFIEKEAPKFINIDFQTKTLIAQNIQEATSGCFTTAQKRV 177

Qy  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPT 210
         || ||||||||  |  ||:  ||
Db  178 YSLMENNSYPRFLSSEFYQDLCK-KPQITT 206

RESULT 6
US-08-748-483-1
; Sequence 1, Application US/08748483
; Patent No. 5955314
; GENERAL INFORMATION:
;     APPLICANT: Hillman, Jennifer L.
;     APPLICANT: Goli, Surya K.
;     TITLE OF INVENTION: A NOVEL REGULATOR OF CELL SIGNALLING
;     NUMBER OF SEQUENCES: 5
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;       STREET: 3174 Porter Drive
;       CITY: Palo Alto
;       STATE: CA
;       COUNTRY: US
;       ZIP: 94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette
;       COMPUTER: IBM Compatible
;       OPERATING SYSTEM: DOS
;       SOFTWARE: FastSEQ Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/748,483
;       FILING DATE: Herewith
;       CLASSIFICATION: 530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Billings, Lucy J.
;       REGISTRATION NUMBER: 36,749
;       REFERENCE/DOCKET NUMBER: PF-0157 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 415-855-0555
;       TELEFAX: 415-845-4166
```

```
;       TELEX:
; INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 181 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     IMMEDIATE SOURCE:
;       LIBRARY: Consensus
;       CLONE: 57362
US-08-748-483-1

Query Match         35.2%;  Score 427.5; DB 2; Length 181;
Best Local Similarity 50.3%;  Pred. No. 1.1e-35;
Matches   84;  Conservative   33; Mismatches   47;  Indels    3;  Gaps    1;

Qy   40 RAKEKRNRLSLLLQRPDFHGE--TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDG 96
         ||||  ||  ||:| ||:|  : |  :| ||||||||
Db   15 RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKJLQNNYG 74

Qy   97 VDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFH 156
         ||: :: | :|||||  : |||| ||||::| ||
Db   75 LASFKSFLKSEFSSENLEFWIACEDYKKIKSPAKMAEKAKQIYEEFIQTEAPKEVNIDHF 134

Qy  157 TKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIE 203
         ||: |:|| :|| | :|| ||||||| :|:  : |
Db  135 TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK 181

RESULT 7
US-08-829-110-6
; Sequence 6, Application US/08829110
; Patent No. 5882890
; GENERAL INFORMATION:
;     APPLICANT: Hillman, Jennifer L.
;     APPLICANT: Goli, Surya K.
;     TITLE OF INVENTION: NOVEL REGULATORS OF G-PROTEIN
;     TITLE OF INVENTION: SIGNALING
;     NUMBER OF SEQUENCES: 6
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Incyte Pharmaceuticals, Inc.
;       STREET: 3174 Porter Drive
;       CITY: Palo Alto
;       STATE: CA
;       COUNTRY: USA
;       ZIP: 94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette
;       COMPUTER: IBM Compatible
;       OPERATING SYSTEM: DOS
;       SOFTWARE: FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/829,110
;       FILING DATE: Filed Herewith
;       CLASSIFICATION: 514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Billings, Lucy J.
;       REGISTRATION NUMBER: 36,749
;       REFERENCE/DOCKET NUMBER: PF-0259 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 415-855-0555
;       TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 205 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     IMMEDIATE SOURCE:
;       LIBRARY: GenBank
```

```
;
US-08-829-110-6

Query Match             32.7%;    Score 397.5;  DB 2;  Length 205;
Best Local Similarity   47.0%;    Pred. No. 1.4e-32;
Matches    77;  Conservative    35;  Mismatches   49;  Indels    3;  Gaps    2;

Qy       41  AKEKRNRLSLLQRPDF--HGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVD 98
             |||::||||||||     ::    ::.::::||||||   :||:
Db       16  AKDMKHRLGFLLQKSDSCEHNSSHNKKDKVVICQ-RVSQEEVKKWAESLENLISHECGLA 74

Qy       99  AFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTK 158
             |||||||||||:||||||||:|::|||::|||| ::|  ::|||||||||||||:|
Db       75  AFKAFLKSEYSEENIDFWISCEEYKKIKSPSKLSPKAKKIYNEFISVQATKEVNLDSCTR 134

Qy      159  EVIAKSIAQPTLHSFDTAQSRVQLMEHDSYKRFLKSETYLHLI 202
             ::|||:||:  :|||     ||:||||||||||:|
Db      135  EETSRNMLEPTITCFDEAQKKIFNLMEKDSYRRFLKSRFYLDLV 178

RESULT 8
US-08-748-483-5
; Sequence 5, Application US/08748483
; Patent No. 5955314
; GENERAL INFORMATION:
;  APPLICANT: Hillman, Jennifer L.
;  APPLICANT: Goli, Surya K.
;  TITLE OF INVENTION: A NOVEL REGULATOR OF CELL SIGNALLING
;  NUMBER OF SEQUENCES: 5
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 Porter Drive
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: US
;   ZIP: 94304
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ Version 2.0
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/748,483
;   FILING DATE: Herewith
;   CLASSIFICATION: 530
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;  ATTORNEY/AGENT INFORMATION:
;   NAME: Billings, Lucy J.
;   REGISTRATION NUMBER: 36,749
;   REFERENCE/DOCKET NUMBER: PF-0157 US
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-855-0555
;   TELEFAX: 415-845-4166
;   TELEX:
; INFORMATION FOR SEQ ID NO: 5:
;  SEQUENCE CHARACTERISTICS:
;   LENGTH: 205 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;  IMMEDIATE SOURCE:
;   LIBRARY: GenBank
;   CLONE: 1216373
US-08-748-483-5

Query Match             32.7%;    Score 397.5;  DB 2;  Length 205;
Best Local Similarity   47.0%;    Pred. No. 1.4e-32;
Matches    77;  Conservative    35;  Mismatches   49;  Indels    3;  Gaps    2;

Qy       41  AKEKRNRLSLLQRPDF--HGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVD 98
             |||::||||||||     ::    ::.::::||||||   :||:
```

```
Db       16  AKDMKHRLGFLLQKSDSCEHNSSHNKKDKVVICQ-RVSQEEVKKWAESLENLISHECGLA 74

Qy       99  AFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTK 158
             |||||||||||:||||||||:|::|||::|||| ::|  ::|||||||||||||:|
Db       75  AFKAFLKSEYSEENIDFWISCEEYKKIKSPSKLSPKAKKIYNEFISVQATKEVNLDSCTR 134

Qy      159  EVIAKSIAQPTLHSFDTAQSRVQLMEHDSYKRFLKSETYLHLI 202
             ::|||:||:  :|||     ||:||||||||||:|
Db      135  EETSRNMLEPTITCFDEAQKKIFNLMEKDSYRRFLKSRFYLDLV 178

RESULT 9
US-08-890-865A-11
; Sequence 11, Application US/08890865A
; Patent No. 6307019
; GENERAL INFORMATION:
;  APPLICANT: Constantini, Franklin
;  APPLICANT: Zeng, Li
;  TITLE OF INVENTION: AXIN GENE AND USES THEREOF
;  NUMBER OF SEQUENCES: 23
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Cooper & Dunham LLP
;   STREET: 1185 Avenue of the Americas
;   CITY: New York
;   STATE: New York
;   COUNTRY: US
;   ZIP: 10036
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/890,865A
;   FILING DATE: 10-JUL-1997
;   CLASSIFICATION: 435
;  ATTORNEY/AGENT INFORMATION:
;   NAME: White, John P
;   REGISTRATION NUMBER: 28,678
;   REFERENCE/DOCKET NUMBER: 0575/54249
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212)278-0400
;   TELEFAX: (212)391-0526
; INFORMATION FOR SEQ ID NO: 11:
;  SEQUENCE CHARACTERISTICS:
;   LENGTH: 120 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;  MOLECULE TYPE: Protein
US-08-890-865A-11

Query Match             30.5%;    Score 370;  DB 4;  Length 120;
Best Local Similarity   58.8%;    Pred. No. 4.3e-30;
Matches    70;  Conservative    21;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       83  WAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKF 142
             |:||||||||||:     |||||||||:|||||||||:||||:|||:||::||||||||
Db        2  WSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACEDFKKTKSPQKLSSKARKIYTDF 61

Qy      143  IQNDAPKEVNIDFHTKEVIAKSIAQPTLHSPDTAQSRVYQLMEHDSYKRFLKSETYLHL 201
             |::||||||||:|:|||||||::||||  ||  ||:||||::|||||||:|
Db       62  IEKEAPKEINIDFQTKTLIAGNIQEATSGCFTTAQKRVYSLMENNSYPRFLESEFYQDL 120

RESULT 10
US-08-274-318-2
; Sequence 2, Application US/08274318
; Patent No. 5667987
; GENERAL INFORMATION:
;  APPLICANT: Buckbinder, Leonard
;  APPLICANT: Talbott, Randy
;  APPLICANT: Seizinger, Bernd R.
```

```
APPLICANT: Kley, Nikolai
TITLE OF INVENTION: No. 5667987el p53 Response Genes
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Bristol-Myers Squibb Company
  STREET: 3005 First Avenue
  CITY: Seattle
  STATE: Washington
  COUNTRY: USA
  ZIP: 98121
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/274,318
  FILING DATE: 11-JUL-1994
  CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
  NAME: Bogden, James M.
  REGISTRATION NUMBER: 32,962
  REFERENCE/DOCKET NUMBER: ON0127-
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 206-727-3688
  TELEFAX: 206-727-3601
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
  LENGTH: 202 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-274-318-2

Query Match     30.0%;  Score 364;  DB 1;  Length 202;
Best Local Similarity   45.8%;  Pred. No. 3.5e-29;
Matches  76;  Conservative  29;  Mismatches  57;  Indels  4;  Gaps  1;

Qy    40 RAKEKRNRLSLLQRPDFHGET---QASRSALLAKETRVSPEEAVKWAESFDKLLSHRD 95
            |||||| ||| :: | |    :.:.:    ::  ::: | :: : || ||::|||||
Db    15 RAKEFKTRLGIFLHKSELGCDTGSTGKSEWGSKSHSKENRNFSEDVLGWRESFDLLLSSKN 74

Qy    96 GVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDF 155
            ||.||  ||||||||:|||:||.||:|||  :  | |  :.:||||:||:: ||||||||
Db    75 GVAAFHAFLKTEFSEENLEFWLACEEFKKIRSATKLASRAHQIFEEFICSEAPKEVNIDH 134

Qy   156 HTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHL 201
            :|::| ::  |.: |  :: .||::|:|||||||:|||||:: :|
Db   135 ETRELTRMNLQTATATCFDAAQGKTRTLMEKDSYPRFLKSPAYRDL 180

RESULT 11
US-08-754-108-2
; Sequence 2, Application US/08754108
; Patent No. 5886149
GENERAL INFORMATION:
  APPLICANT: Buckbinder, Leonard
  APPLICANT: Talbott, Randy
  APPLICANT: Seizinger, Bernd R.
  APPLICANT: Kley, Nikolai
TITLE OF INVENTION: No. 5886149el p53 Response Genes
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Bristol-Myers Squibb Company
  STREET: 3005 First Avenue
  CITY: Seattle
  STATE: Washington
  COUNTRY: USA
  ZIP: 98121
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/754,108
  FILING DATE:
  CLASSIFICATION: 530
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/274,318
  FILING DATE: 11-JUL-1994
ATTORNEY/AGENT INFORMATION:
  NAME: Bogden, James M.
  REGISTRATION NUMBER: 32,962
  REFERENCE/DOCKET NUMBER: ON0127-
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 206-727-3688
  TELEFAX: 206-727-3601
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
  LENGTH: 202 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-754-108-2

Query Match     30.0%;  Score 364;  DB 2;  Length 202;
Best Local Similarity   45.8%;  Pred. No. 3.5e-29;
Matches  76;  Conservative  29;  Mismatches  57;  Indels  4;  Gaps  1;

Qy    40 RAKEKRNRLSLLQRPDFHGET---QASRSALLAKETRVSPEEAVKWAESFDKLLSHRD 95
            |||||| ||| :: | |    :.:.:    ::  ::: | :: : || ||::|||||
Db    15 RAKEFKTRLGIFLHKSELGCDTGSTGKSEWGSKSHSKENRNFSEDVLGWRESFDLLLSSKN 74

Qy    96 GVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDF 155
            ||.||  ||||||||:|||:||.||:|||  :  | |  :.:||||:||:: ||||||||
Db    75 GVAAFHAFLKTEFSEENLEFWLACEEFKKIRSATKLASRAHQIFEEFICSEAPKEVNIDH 134

Qy   156 HTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHL 201
            :|::| ::  |.: |  :: .||::|:|||||||:|||||:: :|
Db   135 ETRELTRMNLQTATATCFDAAQGKTRTLMEKDSYPRFLKSPAYRDL 180

RESULT 12
US-08-463-081B-2
; Sequence 2, Application US/08463081B
; Patent No. 5871960 5837487
; Patent No. 5871960 5837487
GENERAL INFORMATION:
  APPLICANT: Smith, Kendall A. & Beadling, Carol
TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
TITLE OF INVENTION: Vector and Transformed Cell Thereof, and Expression Thereof
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
  ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
  STREET: 444 South Flower St. - Suite 1900
  CITY: Los Angeles
  STATE: California
  COUNTRY: USA
  ZIP: 90071
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0,
  SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/463,081B
  FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/104,736
  FILING DATE: 10-AUG-1993
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/796,066
  FILING DATE: 20-NOV-91
ATTORNEY/AGENT INFORMATION:
```

us-09-894-749-4.rai

;   MOLECULE TYPE: protein
US-08-461-379A-2

Query Match   29.9%;   Score 363; DB 2;   Length 202;
Best Local Similarity   45.8%;   Pred. No. 4.5e-29;
Matches   76; Conservative   29; Mismatches   57; Indels   4; Gaps   1;

Qy   40 RAKEKRNRLSLLLQRPDFHGETQASR---SALLAKETRVSPEEAVKWAESFDKLLSHRD 95
            ||||-|||:||||:|||:||::||| ::|| ||| |||||||||| :||||||| ::
Db   15 RAKEFKTRLGIFLHKSELGCDTGSTGKFEWGSKHSKENRNFSEDVLGWRESFDLLLSSKN 74

Qy   96 GVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDF 155
            ||||||:||||||||||:||::|| : || |:|:|:| :|| :||::|:||::|||||:|
Db   75 GVAAFHAFLKTEFSEENLEFWLACEEFFKKIRSATKLLASRAHQIFEEFICSEAPKEVNIDH 134

Qy   156 HTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHL 201
            |:|:| :|| ::|||| ||| : | ||| :| |||||||
Db   135 ETRELTRMNLQTATATCFDAAQGKTRTLMEKDSYPRFLKSPAYRDL 180

RESULT 13
US-08-461-379A-2
; Sequence 2, Application US/08461379A
; Patent No. 5871961
; GENERAL INFORMATION:
;   APPLICANT: Smith, Kendall A. & Beadling, Carol
;   TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
;   TITLE OF INVENTION: Vector and Transformed Cell Thereof, and
;   TITLE OF INVENTION: Expression Thereof
;   NUMBER OF SEQUENCES: 35
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Ratner & Prestia
;   ADDRESSEE: (B) STREET:One Westlakes-Berwyn
;   CITY: Valley Forge
;   STATE: Pennsylvania
;   COUNTRY: USA
;   ZIP: 19482
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0,
;   SOFTWARE: Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/461,379A
;   FILING DATE: 5-JUNE-1995
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: USSN 08/330,108; 08/104,736
;   APPLICATION NUMBER: & 07/796,066
;   FILING DATE: 27-OCT-1994; 10-AUG-1993 & 20-NOV-91
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Viviana Amzel, Ph. D.
;   REGISTRATION NUMBER: 30,930
;   REFERENCE/DOCKET NUMBER: DART-070
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (610)470-0700
;   TELEFAX: (610)470-0701
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 202 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear

;   NAME: Viviana Amzel, Ph. D.
;   REGISTRATION NUMBER: 30,930
;   REFERENCE/DOCKET NUMBER: P66 38150 (DART-060)
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (213) 622-7700
;   TELEFAX: (213) 489-4210
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 202 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-463-081B-2

Query Match   29.9%;   Score 363; DB 2;   Length 202;
Best Local Similarity   45.8%;   Pred. No. 4.5e-29;
Matches   76; Conservative   29; Mismatches   57; Indels   4; Gaps   1;

Qy   40 RAKEKRNRLSLLLQRPDFHGETQASR---SALLAKETRVSPEEAVKWAESFDKLLSHRD 95
            ||||-|||:||||:|||:||::||| ::|| ||| |||||||||| :||||||| ::
Db   15 RAKEFKTRLGIFLHKSELGCDTGSTGKFEWGSKHSKENRNFSEDVLGWRESFDLLLSSKN 74

Qy   96 GVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDF 155
            ||||||:||||||||||:||::|| : || |:|:|:| :|| :||::|:||::|||||:|
Db   75 GVAAFHAFLKTEFSEENLEFWLACEEFFKKIRSATKLLASRAHQIFEEFICSEAPKEVNIDH 134

Qy   156 HTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHL 201
            |:|:| :|| ::|||| ||| : | ||| :| |||||||
Db   135 ETRELTRMNLQTATATCFDAAQGKTRTLMEKDSYPRFLKSPAYRDL 180

RESULT 14
US-08-462-390B-2
; Sequence 2, Application US/08462390B
; Patent No. 5882894
; GENERAL INFORMATION:
;   APPLICANT: Smith, K. A. & Beadling, C.
;   TITLE OF INVENTION: Nucleic Acids Encoding CR8 Polypeptide, Vector and
;   TITLE OF INVENTION: Transformed Cell Thereof, and Expression Thereof
;   NUMBER OF SEQUENCES: 35
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Ratner & Prestia
;   ADDRESSEE: (B) STREET:One Westlakes-Berwyn
;   CITY: Valley Forge
;   STATE: Pennsylvania
;   COUNTRY: USA
;   ZIP: 19482
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/462,390B
;   FILING DATE: 5-JUNE-1995
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: USSN 08/330,108
;   FILING DATE: 27-OCT-1994
;   APPLICATION NUMBER: USSN 08/104,736
;   FILING DATE: 10-AUG-1993
;   APPLICATION NUMBER: USSN 07/796,066
;   FILING DATE: 20-NOV-91
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Viviana Amzel, Ph. D.
;   REGISTRATION NUMBER: 30,930
;   REFERENCE/DOCKET NUMBER: DART-040
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (610)407-0700
;   TELEFAX: (610)407-0701
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 202 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-462-390B-2

Query Match   29.9%;   Score 363; DB 2;   Length 202;
Best Local Similarity   45.8%;   Pred. No. 4.5e-29;
Matches   76; Conservative   29; Mismatches   57; Indels   4; Gaps   1;

Qy   40 RAKEKRNRLSLLLQRPDFHGETQASR---SALLAKETRVSPEEAVKWAESFDKLLSHRD 95
            ||||-|||:||||:|||:||::||| ::|| ||| |||||||||| :||||||| ::
Db   15 RAKEFKTRLGIFLHKSELGCDTGSTGKFEWGSKHSKENRNFSEDVLGWRESFDLLLSSKN 74

```
Qy     96  GVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDF  155
           || |||||||||||||  ||| :  : | :|:||  ::||||| :|||||||
Db     75  GVAAFHAFLKTEFSEENLEFWLACEEFKKIRSATKLASRAHQIFEEFICSEAPKEVNIDH  134

Qy    156  HTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHL  201
           |:|: :: ||||| :||| ||||  |||| || ||||
Db    135  ETRELTRMNLQTATATCFDAAQGKTRTLMEKDSYPRFLKSPAYRDL  180


RESULT 15
US-08-870-815-4
; Sequence 4, Application US/08870815
; Patent No. 6020142
; GENERAL INFORMATION:
;  APPLICANT: Levinson, Douglas A.
;  APPLICANT: Gimeno, Carlos J.
;  TITLE OF INVENTION:  RATH GENES AND POLYPEPTIDES AND METHODS
;  TITLE OF INVENTION:  FOR THE TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
;  NUMBER OF SEQUENCES:  15
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  Pennie & Edmonds LLP
;   STREET:  1155 Avenue of the Americas
;   CITY:  New York
;   STATE:  New York
;   COUNTRY:  USA
;   ZIP:  10036-2711
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/870,815
;   FILING DATE:  06-JUN-1997
;   CLASSIFICATION:  435
;  ATTORNEY/AGENT INFORMATION:
;   NAME:  Laura A. Coruzzi
;   REGISTRATION NUMBER:  30,742
;   REFERENCE/DOCKET NUMBER:  7853-074
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (212) 790-9090
;   TELEFAX:  (212) 896-8864/9741
;   TELEX:  66141 PENNIE
; INFORMATION FOR SEQ ID NO:  4:
;  SEQUENCE CHARACTERISTICS:
;   LENGTH:  202 amino acids
;   TYPE:  amino acid
;   STRANDEDNESS:
;   TOPOLOGY:  unknown
;  MOLECULE TYPE:  protein
US-08-870-815-4

Query Match             29.9%;  Score 363;  DB 3;  Length 202;
Best Local Similarity   45.8%;  Pred. No. 4.5e-29;
Matches   76;  Conservative   29;  Mismatches   57;  Indels    4;  Gaps    1;

Qy     40  RAKEKRNRLSLLLQRPDFHGETQASR---SALLAKETRVSPBEAVKWAESFDKLLSHRD  95
           ||||  |||  :  : ::    : :|: :| |  ||  : :|||||| ::
Db     15  RAKEFKTRLGIFLHKSELGCDTGSTGKFEWGSKHSKENRNFSEDVLGWRESFDLLJSKN  74

Qy     96  GVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDF  155
           || |||||||||||||  ||| :  : | :|:||  ::||||| :|||||||
Db     75  GVAAFHAFLKTEFSEENLEFWLACEEFKKIRSATKLASRAHQIFEEFICSEAPKEVNIDH  134

Qy    156  HTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHL  201
           |:|: :: ||||| :||| ||||  |||| || ||||
Db    135  ETRELTRMNLQTATATCFDAAQGKTRTLMEKDSYPRFLKSPAYRDL  180


Search completed: August 20, 2004, 16:48:33
Job time : 19.5 secs
```

OM protein - protein search, using sw model

```
                    Copyright   GenCore version 5.1.6
                    Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:       August 20, 2004, 16:47:08 ; Search time 239 Seconds
                                           (without alignments)
                                           308.996 Million cell updates/sec

Title:        US-09-894-749-4
Perfect score: 1214
Sequence:     1 MDMSLVFFSQLNMCESKEKT........SRSFTYNDFQDVKSDVAIWL 235
Scoring table: BLOSUM62
              Gapop 10.0 , Gapext 0.5
Searched:     1295152 seqs, 31255058 residues

Total number of hits satisfying chosen parameters:     1295152

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Published Applications AA:*
              1:  /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
              2:  /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
              3:  /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
              4:  /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
              5:  /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
              6:  /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
              7:  /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
              8:  /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
              9:  /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
              10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
              11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
              12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
              13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
              14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
              15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
              16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
              17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
              18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1214	100.0	235	9	US-09-894-749-4	Sequence 4, Appli
2	1032	85.0	235	9	US-09-894-749-2	Sequence 2, Appli
3	1032	85.0	235	12	US-10-258-371B-20	Sequence 20, Appl
4	994	81.9	227	9	US-09-867-550-848	Sequence 848, App
5	429.5	35.4	211	9	US-09-206-639-4	Sequence 4, Appli
6	429.5	35.4	211	12	US-10-258-371B-24	Sequence 24, Appl
7	429.5	35.4	211	16	US-10-408-765A-493	Sequence 493, App
8	429.5	35.4	220	9	US-09-925-300-1507	Sequence 1507, Ap
9	427.5	35.2	181	9	US-09-206-639-1	Sequence 1, Appli
10	427.5	35.2	181	12	US-10-258-371B-22	Sequence 22, Appl
11	427.5	35.2	190	14	US-10-411-224-197	Sequence 197, App
12	427.5	35.2	190	15	US-10-047-021-197	Sequence 197, App
13	412.5	34.0	181	14	US-10-398-953-1	Sequence 1, Appli
14	407.5	33.6	180	12	US-10-258-371B-27	Sequence 27, Appl
15	407.5	33.6	180	14	US-10-275-555-2	Sequence 2, Appli
16	397.5	32.7	205	9	US-09-206-639-5	Sequence 5, Appli
17	397.5	32.7	205	10	US-09-939-209A-2	Sequence 2, Appli
18	397.5	32.7	205	12	US-10-258-371B-21	Sequence 21, Appl
19	396	32.6	930	14	US-10-113-794A-1	Sequence 1, Appli
20	383	31.5	519	12	US-10-258-371B-28	Sequence 28, Appl
21	383	31.5	519	14	US-10-113-794A-2	Sequence 2, Appli
22	383	31.5	519	14	US-10-428-487-14	Sequence 14, Appl
23	383	31.5	591	15	US-10-108-260A-3970	Sequence 3970, Ap
24	381	31.4	776	12	US-10-087-192-1728	Sequence 1728, Ap
25	380	31.3	284	15	US-10-094-749-1650	Sequence 1650, Ap
26	375.5	30.9	923	12	US-10-114-270-152	Sequence 152, App
27	364	30.0	202	14	US-10-171-311-202	Sequence 202, App
28	363	29.9	202	9	US-09-873-438-4	Sequence 4, Appli
29	363	29.9	204	12	US-10-258-371B-23	Sequence 23, Appl
30	362	29.8	81	12	US-10-258-371B-12	Sequence 12, Appl
31	360	29.7	128	15	US-10-352-843-1	Sequence 1, Appli
32	360	29.7	128	15	US-10-352-843-3	Sequence 3, Appli
33	360	29.7	167	10	US-09-942-055A-1	Sequence 1, Appli
34	348	28.7	201	9	US-09-873-438-2	Sequence 2, Appli
35	347.5	28.6	196	9	US-09-206-639-3	Sequence 3, Appli
36	347.5	28.6	196	9	US-09-736-457-339	Sequence 339, App
37	347.5	28.6	196	9	US-09-902-941-339	Sequence 339, App
38	347.5	28.6	196	9	US-09-849-626-339	Sequence 339, App
39	347.5	28.6	196	10	US-09-476-300-339	Sequence 339, App
40	347.5	28.6	196	12	US-10-283-017-339	Sequence 339, App
41	347.5	28.6	196	14	US-10-017-754-339	Sequence 339, App
42	347.5	28.6	196	14	US-10-113-872-339	Sequence 339, App
43	347.5	28.6	196	14	US-10-247-671-176	Sequence 176, App
44	347.5	28.6	217	9	US-09-925-301-1292	Sequence 1292, Ap
45	345.5	28.5	196	12	US-10-170-385-375	Sequence 375, App

ALIGNMENTS

```
RESULT 1
US-09-894-749-4
; Sequence 4, Application US/0894749
; Patent No. US2002008168A1
; GENERAL INFORMATION:
;  APPLICANT: Hodge, Martin R.
;  APPLICANT: Yowe, David
;  TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
;  FILE REFERENCE: 5800-19, 035800/174680
;  CURRENT APPLICATION NUMBER: US/09/894,749
;  CURRENT FILING DATE: 2001-06-27
;  PRIOR APPLICATION NUMBER: 09/244,314
;  PRIOR FILING DATE: 1999-02-04
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;  LENGTH: 235
;  TYPE: PRT
;  ORGANISM: Mus sp.
US-09-894-749-4

  Query Match          100.0%;  Score 1214;  DB 9;  Length 235;
  Best Local Similarity 100.0%;  Pred. No. 7.4e-110;
  Matches   235;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1  MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE  60

Qy    61  TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE  120

Qy   121  DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV  180

Qy   181  YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL  235
```

```
Db  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 2
US-09-894-749-2
; Sequence 2, Application US/09894749
; Patent No. US20020081683A1
; GENERAL INFORMATION:
;　APPLICANT: Hodge, Martin R.
;　APPLICANT: Ycwe, David
;　TITLE OF INVENTION: RGS-Containing Molecules and Uses Thereof
;　FILE REFERENCE: 5800-19　035800/174680
;　CURRENT APPLICATION NUMBER: US/09/894,749
;　CURRENT FILING DATE: 2001-06-27
;　PRIOR APPLICATION NUMBER: 09/244,314
;　PRIOR FILING DATE: 1999-02-04
;　NUMBER OF SEQ ID NOS: 4
;　SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 2
;　LENGTH: 235
;　TYPE: PRT
;　ORGANISM: Homo sapiens
US-09-894-749-2

Query Match　　　　85.0%;　Score 1032;　DB 9;　Length 235;
Best Local Similarity　83.8%;　Pred. No. 3.9e-92;
Matches 197; Conservative 19; Mismatches 19; Indels　0; Gaps　0;

```
Qy    1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLQRPDFHGE 60
          |:.:||:.||||||||||||||:||:||||:||:.:.||:.||:|
Db    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy   61 TQASRSALLAKETRVSPEEAVKWAASFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
          |.:||| ||||||||||||||||||||||||||||:||
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy  121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
          |||||| ||:||||||||||||||::|||||||||:||
Db  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
          ||||:|||||||||||||||||||||||||||||||:||||||||||||||||
Db  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
```

RESULT 3
US-10-258-371B-20
; Sequence 20, Application US/10258371B
; Publication No. US20040067903A1
; GENERAL INFORMATION:
;　APPLICANT: WILLIAMS-GAGNON, Alison
;　APPLICANT: MURRAY, David L
;　TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING,
;　TITLE OF INVENTION: RGS18, AND USES THEREOF
;　FILE REFERENCE: A3656 US PCT
;　CURRENT APPLICATION NUMBER: US/10/258,371B
;　CURRENT FILING DATE: 2003-06-04
;　PRIOR APPLICATION NUMBER: GB001883.334
;　PRIOR FILING DATE: 2000-08-02
;　PRIOR APPLICATION NUMBER: US60/200,786
;　PRIOR FILING DATE: 2000-04-28
;　NUMBER OF SEQ ID NOS: 38
;　SOFTWARE: PatentIn version 3.2
;
; SEQ ID NO 20
;　LENGTH: 235
;　TYPE: PRT
;　ORGANISM: Homo sapiens
US-10-258-371B-20

Query Match　　　　85.0%;　Score 1032;　DB 12;　Length 235;
Best Local Similarity　83.8%;　Pred. No. 3.9e-92;
Matches 197; Conservative 19; Mismatches 19; Indels　0; Gaps　0;

```
Qy    1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLQRPDFHGE 60
          |:.:||:.||||||||||||||:||:||||:||:.:.||:.||:|
Db    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy   61 TQASRSALLAKETRVSPEEAVKWAASFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
          |.:||| ||||||||||||||||||||||||||||:||
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy  121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
          |||||| ||:||||||||||||||::|||||||||:||
Db  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
          ||||:|||||||||||||||||||||||||||||||:||||||||||||||||
Db  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDVQSDVAIWL 235
```

RESULT 4
US-09-867-550-848
; Sequence 848, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
;　APPLICANT: Leach, Martin D.
;　APPLICANT: Mehraban, Fuad.
;　APPLICANT: Conley, Pamela
;　APPLICANT: Law, Debbie
;　APPLICANT: Topper, James
;　TITLE OF INVENTION: No. US20020082206A1e1 Polynucleotides from Atherogenic Cells and L
;　TITLE OF INVENTION: Thereby
;　FILE REFERENCE: 21402-013 (Cura-313)
;　CURRENT APPLICATION NUMBER: US/09/867,550
;　CURRENT FILING DATE:　2001-09-20
;　PRIOR APPLICATION NUMBER: USSN 60/208,427
;　PRIOR FILING DATE: 2000-05-30
;　NUMBER OF SEQ ID NOS: 2125
;　SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 848
;　LENGTH: 227
;　TYPE: PRT
;　ORGANISM: Homo sapiens
US-09-867-550-848

Query Match　　　　81.9%;　Score 994;　DB 9;　Length 227;
Best Local Similarity　83.7%;　Pred. No. 1.9e-88;
Matches 190; Conservative 18; Mismatches 19; Indels　0; Gaps　0;

```
Qy    1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLQRPDFHGE 60
          |:.:||:.||||||||||||||:||:||||:||:.:.||:.||:|
Db    1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy   61 TQASRSALLAKETRVSPEEAVKWAASFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
          |.:||| ||||||||||||||||||||||||||||:||
Db   61 TRSSRSGHLAKETRVSPEEAVKWGESFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy  121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
          |||||| ||:||||||||||||||::|||||||||:||
Db  121 DFKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVITNSITQPTLHSFDAAQSRV 180

Qy  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDV 227
          ||||:|||||||||||||||||||||||||||||||:||||||
Db  181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSFTCNEFQDV 227
```

RESULT 5
US-09-206-639-4
; Sequence 4, Application US/09206639
; Patent No. US20020034777A1
; GENERAL INFORMATION:
;　APPLICANT: Hillman, Jennifer L.
;　　　　　　　Goli, Surya K.
;　TITLE OF INVENTION: A NOVEL REGULATOR OF CELL SIGNALLING
;　NUMBER OF SEQUENCES: 5

```
CORRESPONDENCE ADDRESS:
    ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
    STREET: 3174 Porter Drive
    CITY: Palo Alto
    STATE: CA
    COUNTRY: US
    ZIP: 94304
COMPUTER READABLE FORM:
    MEDIUM TYPE: Diskette
    COMPUTER: IBM Compatible
    OPERATING SYSTEM: DOS
    SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/206,639
    FILING DATE: 07-Dec-1998
    CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/748,483
    FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
    NAME: Billings, Lucy J.
    REGISTRATION NUMBER: 36,749
    REFERENCE/DOCKET NUMBER: PF-0157 US
TELECOMMUNICATION INFORMATION:
    TELEPHONE: 415-855-0555
    TELEFAX: 415-845-4166
    TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 4:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 211 amino acids
        TYPE: amino acid
        STRANDEDNESS: single
        TOPOLOGY: linear
    IMMEDIATE SOURCE:
        LIBRARY: GenBank
        CLONE: 292037
    SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-206-639-4

Query Match    35.4%;    Score 429.5;    DB 9;    Length 211;
Best Local Similarity    44.3%;    Pred. No. 1.6e-33;
Matches   93;    Conservative   38;    Mismatches   70;    Indels    9;    Gaps    3;

Qy    4  SLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIR---AKEKRNRLSLLLQRPDFHGE   60
         :|:|: ::  ||:|         |||  ||      |::  :||: ||| |:||:
Db    3  SAMFLAVQHDCRPMDKS-----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSSTPGK   57

Qy   61  TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE  120
           : :  || :|:|||||| ||:|||| :|||||:||| : |||||:|:||||||:|||
Db   58  PKTGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACE  117

Qy  121  DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV  180
         ||||  | |:|: :|::||| ||::::|||:|||| || | : |:||  :|| ||| :||
Db  118  DFKKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRV  177

Qy  181  YQLMEHDSYKRFLKSETYLHLIEGRPQRPT  210
         | ||| :||:||  | | : | ::: ||
Db  178  YSLMENNSYPRFLESEFYQDLCK-KPQITT  206
```

RESULT 6
US-10-258-371B-24
```
; Sequence 24, Application US/10258371B
; Publication No. US2004067903A1
; GENERAL INFORMATION:
    APPLICANT: WILLIAMS-GAGNON, Alison
    APPLICANT: MURRAY, David L
    TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING,
    TITLE OF INVENTION: AND USES THEREOF
    FILE REFERENCE: A3656 US PCT
    CURRENT APPLICATION NUMBER: US/10/258,371B
    CURRENT FILING DATE: 2003-06-04
    PRIOR APPLICATION NUMBER: GB001883.334
    PRIOR FILING DATE: 2000-08-02
    PRIOR APPLICATION NUMBER: US60/200,786
    PRIOR FILING DATE: 2000-04-28
    NUMBER OF SEQ ID NOS: 38
    SOFTWARE: PatentIn version 3.2
SEQ ID NO 24
    LENGTH: 211
    TYPE: PRT
    ORGANISM: Homo sapiens
US-10-258-371B-24

Query Match    35.4%;    Score 429.5;    DB 12;    Length 211;
Best Local Similarity    44.3%;    Pred. No. 1.6e-33;
Matches   93;    Conservative   38;    Mismatches   70;    Indels    9;    Gaps    3;

Qy    4  SLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIR---AKEKRNRLSLLLQRPDFHGE   60
         :|:|: ::  ||:|         |||  ||      |::  :||: ||| |:||:
Db    3  SAMFLAVQHDCRPMDKS-----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSSTPGK   57

Qy   61  TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE  120
           : :  || :|:|||||| ||:|||| :|||||:||| : |||||:|:||||||:|||
Db   58  PKTGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACE  117

Qy  121  DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV  180
         ||||  | |:|: :|::||| ||::::|||:|||| || | : |:||  :|| ||| :||
Db  118  DFKKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRV  177

Qy  181  YQLMEHDSYKRFLKSETYLHLIEGRPQRPT  210
         | ||| :||:||  | | : | ::: ||
Db  178  YSLMENNSYPRFLESEFYQDLCK-KPQITT  206
```

RESULT 7
US-10-408-765A-493
```
; Sequence 493, Application US/1040876SA
; Publication No. US20040101874A1
; GENERAL INFORMATION:
    APPLICANT: Ghosh, Sounitra S.
    APPLICANT: Fahy, Eoin D.
    APPLICANT: Zhang, Bing
    APPLICANT: Gibson, Bradford W.
    APPLICANT: Taylor, Steven W.
    APPLICANT: Glenn, Gary M.
    APPLICANT: Warnock, Dale E.
    TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
    TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
    FILE REFERENCE: 660088.465
    CURRENT APPLICATION NUMBER: US/10/408,765A
    CURRENT FILING DATE: 2003-04-04
    NUMBER OF SEQ ID NOS: 3077
    SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 493
    LENGTH: 211
    TYPE: PRT
    ORGANISM: Homo sapiens
US-10-408-765A-493

Query Match    35.4%;    Score 429.5;    DB 16;    Length 211;
Best Local Similarity    44.3%;    Pred. No. 1.6e-33;
Matches   93;    Conservative   38;    Mismatches   70;    Indels    9;    Gaps    3;

Qy    4  SLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIR---AKEKRNRLSLLLQRPDFHGE   60
         :|:|: ::  ||:|         |||  ||      |::  :||: ||| |:||:
Db    3  SAMFLAVQHDCRPMDKS-----AGSGHKSEEKREKMKRTLLKDWKTRLSYFLQNSSTPGK   57

Qy   61  TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE  120
           : :  || :|:|||||| ||:|||| :|||||:||| : |||||:|:||||||:|||
Db   58  PKTGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACE  117

Qy  121  DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV  180
         ||||  | |:|: :|::||| ||::::|||:|||| || | : |:||  :|| ||| :||
Db  118  DFKKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRV  177
```

```
Qy    181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPT 210
          |||||:||:||:|||  : :||:||
Db    178 YSLMENNSYPRFLESEFYQDLCK-KPQITT 206

RESULT 8
US-09-925-300-1507
; Sequence 1507, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
;   APPLICANT: Craig Rosen,
;   APPLICANT: Steve Ruben
;   TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;   FILE REFERENCE: PA101
;   CURRENT APPLICATION NUMBER: US/09/925,300
;   CURRENT FILING DATE: 2001-08-10
;   PRIOR APPLICATION NUMBER: PCT/US00/05988
;   PRIOR FILING DATE: 2000-03-08
;   PRIOR APPLICATION NUMBER: 60/124,270
;   PRIOR FILING DATE: 1999-03-12
;   NUMBER OF SEQ ID NOS: 1890
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1507
;   LENGTH: 220
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-300-1507

Query Match        35.4%;  Score 429.5; DB 9;  Length 220;
Best Local Similarity 44.3%;  Pred. No. 1.7e-33;
Matches. 93;  Conservative 38;  Mismatches 70;  Indels 9;  Gaps 3;

Qy    4 SLVFFSQLNMCESKEKTFPKLMGSGKETSIEAKIR---AKEKRNRLSLLQRPDFHGE 60
          |||:  :         ||: :     |: : ||||||||:||||
Db   12 SAMFLAVQHDCRPMDKS-----AGSGHKSEKREKMKRTLLKDWKTRLSYFLQNSSTPGK 66

Qy   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
          |||||||||||||||||||||||||||||:|||||  :||||||||||||||
Db   67 PKTGKKSKQQAFIKPSPEEAQLWSEAFDLLASKYGLAAFRAFLKSEFCEENIEFWLACE 126

Qy  121 DFKKCCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHIKEVIAKSIAQPTLHSFDTAQSRV 180
          ||||| :||||| ||:|| |||||||||||||||  |||  ||||||||||||:||
Db  127 DFKKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQKTLIAQNIQEATSGCFTTAQKRV 186

Qy  181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPT 210
          |||||:||:||:|||  : :||:||
Db  187 YSLMENNSYPRFLESEFYQDLCK-KPQITT 215

RESULT 9
US-09-206-639-1
; Sequence 1, Application US/09206639
; Patent No. US20020034777A1
; GENERAL INFORMATION:
;   APPLICANT: Hillman, Jennifer L.
;   Goli, Surya K.
;   TITLE OF INVENTION: A NOVEL REGULATOR OF CELL SIGNALLING
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;     STREET: 3174 Porter Drive
;     CITY: Palo Alto
;     STATE: CA
;     COUNTRY: US
;     ZIP: 94304
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FastSEQ Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/206,639
;     FILING DATE: 07-Dec-1998
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/748,483
;     FILING DATE: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Billings, Lucy J.
;     REGISTRATION NUMBER: 36,749
;     REFERENCE/DOCKET NUMBER: PF-0157 US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 415-885-0555
;     TELEFAX: 415-845-4166
;     TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 181 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;     LIBRARY: Consensus
;     CLONE: 57362
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-206-639-1

Query Match        35.2%;  Score 427.5; DB 9;  Length 181;
Best Local Similarity 50.3%;  Pred. No. 2e-33;
Matches 84;  Conservative 33;  Mismatches 47;  Indels 3;  Gaps 1;

Qy   40 RAKEKRNRLSLLQRPDFHGE---TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDG 96
          |||:|||:||:||||  :||   ||||||:||||||||||||| ||:||||||||||
Db   15 RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG 74

Qy   97 VDAFTRFLKTEFSEENIEFWVACEDFKKCCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFH 156
          |||||||||||||||||||||| :||  ||  :||  |:||  :|| ||||  |||||||
Db   75 LASFKSFLKSEFSEENLEFWIACEDYKKKIKSPAKVAEKAKQIYEBFIQTEAPKEVNIDHF 134

Qy  157 TKEVIAKSIAQPTLHSFDTAQSRVQLMEHDSYKRFLKSETYLHLIE 203
          |||  ||||||||||:|||||||:|||:||:|||:|||:||||
Db  135 TKDITMKNLVEFSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK 181

RESULT 10
US-10-258-371B-22
; Sequence 22, Application US/10258371B
; Publication No. US20040067903A1
; GENERAL INFORMATION:
;   APPLICANT: WILLIAMS-GAGNON, Alison
;   APPLICANT: MURRAY, David L
;   TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING,
;   TITLE OF INVENTION: RGS18, AND USES THEREOF
;   FILE REFERENCE: A3656 US PCT
;   CURRENT APPLICATION NUMBER: US/10/258,371B
;   CURRENT FILING DATE: 2003-06-04
;   PRIOR APPLICATION NUMBER: GB001883.334
;   PRIOR FILING DATE: 2000-08-02
;   PRIOR APPLICATION NUMBER: US60/200,786
;   PRIOR FILING DATE: 2000-04-28
;   NUMBER OF SEQ ID NOS: 38
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
;   LENGTH: 181
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-258-371B-22

Query Match        35.2%;  Score 427.5; DB 12;  Length 181;
Best Local Similarity 50.3%;  Pred. No. 2e-33;
Matches 84;  Conservative 33;  Mismatches 47;  Indels 3;  Gaps 1;

Qy   40 RAKEKRNRLSLLQRPDFHGE---TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDG 96
          |||:|||:||:||||  :||   ||||||:||||||||||||| ||:||||||||||
Db   15 RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG 74
```

```
Qy   97  VDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFH  156
         ::|||  ||||:||||:|||:||:||||  ||  ||||:|||||||  :|||||||||||
Db   75  LASFKSFIKSEFSEENLEFWIACEDYKKIKSPAKMAEKAKQIYEEFIQTEAPKEVNIDHF  134

Qy  157  TKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIE  203
         |||||||||||||||||||  :|||||||||||:||||||  :||
Db  135  TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK  181
```

RESULT 11
US-10-411-224-197
; Sequence 197, Application US/10411224
; Publication No. US20030166906A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 50 Human Secreted Proteins
; FILE REFERENCE: P2016P1
; CURRENT FILING DATE: 2003-04-11
; CURRENT APPLICATION NUMBER: US/10/411,224
; PRIOR APPLICATION NUMBER: US/09/722,329
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/262,109
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: 60/057,626
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/057,663
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/057,669
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/058,667
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,974
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,973
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,666
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 197
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-224-197

Query Match   35.2%; Score 427.5; DB 14; Length 190;
Best Local Similarity 50.3%; Pred. No. 2.1e-33;
Matches 84; Conservative 33; Mismatches 47; Indels 3; Gaps 1;

```
Qy   40  RAKEKRNRLSLLQRPDFHGE---TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDG  96
         |||  |||:|| |||  |  :  |  :||  |  |||  :|||||  :|||||  :|||:
Db   24  RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG  83

Qy   97  VDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFH  156
         ::|||  ||||:||||:|||:||:||||  ||  ||||:|||||||  :|||||||||||
Db   84  LASFKSFIKSEFSEENLEFWIACEDYKKIKSPAKMAEKAKQIYEEFIQTEAPKEVNIDHF  143

Qy  157  TKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIE  203
         |||||||||||||||||||  :|||||||||||:||||||  :||
Db  144  TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK  190
```

RESULT 12
US-10-047-021-197
; Sequence 197, Application US/10047021
; Publication No. US20040002591A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 50 Human Secreted Proteins
; FILE REFERENCE: P2016P2
; CURRENT APPLICATION NUMBER: US/10/047,021
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/262,066
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 09/722,329
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/262,109
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: PCT/US98/18360
; PRIOR FILING DATE: 1998-09-03
; PRIOR APPLICATION NUMBER: US 60/057,626
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/057,663
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/057,669
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/058,667
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,974
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,973
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,666
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/090,112
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 197
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-047-021-197

Query Match   35.2%; Score 427.5; DB 15; Length 190;
Best Local Similarity 50.3%; Pred. No. 2.1e-33;
Matches 84; Conservative 33; Mismatches 47; Indels 3; Gaps 1;

```
Qy   40  RAKEKRNRLSLLQRPDFHGE---TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDG  96
         |||  |||:|| |||  |  :  |  :||  |  |||  :|||||  :|||||  :|||:
Db   24  RAKEIKIKLGILLQKPDSVGDLVIPYNEKPEKPAKTQKTSLDEALQWRDSLDKLLQNNYG  83

Qy   97  VDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFH  156
         ::|||  ||||:||||:|||:||:||||  ||  ||||:|||||||  :|||||||||||
Db   84  LASFKSFIKSEFSEENLEFWIACEDYKKIKSPAKMAEKAKQIYEEFIQTEAPKEVNIDHF  143

Qy  157  TKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIE  203
         |||||||||||||||||||  :|||||||||||:||||||  :||
Db  144  TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK  190
```

RESULT 13
US-10-398-953-1
; Sequence 1, Application US/10398953
; Publication No. US20030186304A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: No. US20030186304Alel disease-associated gene and use thereof
; FILE REFERENCE: 2801 USOP
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: JP 2000-319912
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: JP 2000-350183
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Rat
US-10-398-953-1

Query Match   34.0%; Score 412.5; DB 14; Length 181;
Best Local Similarity 49.1%; Pred. No. 5.7e-32;

Matches   86; Conservative   27; Mismatches   41; Indels   21; Gaps   2;

Qy   40 RAKEKRNRLSLLQRPDFHGETQASRSAL---------LAKETRVSPEEAVKWAESF 87
               |||| |: |:|||:|| ||: ||| ::||::|
Db   15 RAKSIKIKIGILLQKPD--------SAVDLVIPYNEKPEKPAKAHKPSLEEVLQWRQSL 65

Qy   88 DKLLSHRDGVDAFTRFLKTEFSSENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDA 147
               ||:||||||||||| :||:|||||:|| ||:||:||:|
Db   66 DKLLQSNYGFASFKSFLKSEFSEENLEFWVACENYKKIKSPIKMAEBKAKQIYEEFIQTEA 125

Qy   148 PKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLI 202
               |||||| ||:: ||: :| :||| ||||||||||||| ||||||||||
Db   126 PKEVNIDHFTKDIITMKNLVEPSPHSFDLAQKRIYALMEKDSLPRFVRSEFYKELI 180

RESULT 14
US-10-258-371B-27
; Sequence 27, Application US/10258371B
; Publication No. US2004006790$A1
; GENERAL INFORMATION:
;   APPLICANT: WILLIAMS-GAGNON, Alison
;   APPLICANT: MURRAY, David L
;   TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A NOVEL REGULATOR OF G PROTEIN SIGNALING,
;   TITLE OF INVENTION: RGS18, AND USES THEREOF
;   FILE REFERENCE: A3656 US PCT
;   CURRENT APPLICATION NUMBER: US/10/258,371B
;   CURRENT FILING DATE: 2003-06-04
;   PRIOR APPLICATION NUMBER: GB001883.334
;   PRIOR FILING DATE: 2000-08-02
;   PRIOR APPLICATION NUMBER: US60/200,786
;   PRIOR FILING DATE: 2000-04-28
;   NUMBER OF SEQ ID NOS: 38
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
;   LENGTH: 180
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
US-10-258-371B-27

Query Match   33.6%;   Score 407.5; DB 12; Length 180;
Best Local Similarity   47.7%;   Pred. No. 1.7e-31;
Matches   83; Conservative   32; Mismatches   50; Indels   9; Gaps   2;

Qy   40 RAKEKRNRLSLLQRPDFHGETQASRSALLAKE-----TRVSPEEAVKWAESFDKLLSHR 94
               |||||||| :| :| ::|| ||||||||:||:|||||||:
Db   9 RNKGMRTRLGCLSHKSD---SCSDFTAILPDKPNRALKRLSTEEATRWADSFDVLLSHK 64

Qy   95 DGVDAFTRFLKTEFSSENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNID 154
               |||||||||||||||||||||||| ::: :|:|:|:|
Db   65 YGVAAFRAFLKTEFSEENLEFWLACEEFKKTRSTAKLVTKAHRIFEEFVDVQAPREVNID 124

Qy   155 FHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEGRPQR 208
               |::| |:: ::|:|:|:||||||||||||||||||||||||| ||: :|
Db   125 FQTREATRKNMQEPSLTCFDQAQGKVHSLMEKDSYPRFLRSKMYLDLLSQSQRR 178

RESULT 15
US-10-275-555-2
; Sequence 2, Application US/10275555
; Publication No. US2003010445$0A1
; GENERAL INFORMATION:
;   APPLICANT: Merck Patent GmbH
;   TITLE OF INVENTION: No. US2003010445$0A1el regulator of G protein signalling (RGS8)
;   FILE REFERENCE: RGS8CWWS
;   CURRENT APPLICATION NUMBER: US/10/275,555
;   CURRENT FILING DATE: 2002-11-07
;   NUMBER OF SEQ ID NOS: 2
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 180
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-275-555-2

Query Match   33.6%;   Score 407.5; DB 14; Length 180;
Best Local Similarity   47.7%;   Pred. No. 1.7e-31;
Matches   83; Conservative   32; Mismatches   50; Indels   9; Gaps   2;

Qy   40 RAKEKRNRLSLLQRPDFHGETQASRSALLAKE-----TRVSPEEAVKWAESFDKLLSHR 94
               |||||||| :| :| ::|| ||||||||:||:|||||||:
Db   9 RNKGMRTRLGCLSHKSD---SCSDFTAILPDKPNRALKRLSTEEATRWADSFDVLLSHK 64

Qy   95 DGVDAFTRFLKTEFSSENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNID 154
               |||||||||||||||||||||||| ::: :|:|:|:|
Db   65 YGVAAFRAFLKTEFSEENLEFWLACEEFKKTRSTAKLVSKAHRIFEEFVDVQAPREVNID 124

Qy   155 FHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEGRPQR 208
               |::| |:: ::|:|:|:||||||||||||||||||||||||| ||: :|
Db   125 FQTREATRKNLQEPSLTCFDQAQGKVHSLMEKDSYPRFLRSKMYLDLLSQSQRR 178

Search completed: August 20, 2004, 17:06:39
Job time : 239 secs

```
              GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:     August 20, 2004, 16:14:26 ; Search time 25 Seconds
                                    (without alignments)
                           904.200 Million cell updates/sec

Title:          US-09-894-749-4
Perfect score:  1214
Sequence:       1 MDMSLVFFSQLNMCESKEKT..........SRSFTYNDFQDVKSDVAIWL 235

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters:       283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     PIR_78:*
             1: pir1:*
             2: pir2:*
             3: pir3:*
             4: pir4:*


Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	% Query Match	Match Length	DB	ID		Description
1	429.5	35.4	211	2	I53020		G-0/G-1 switch reg
2	397.5	32.7	205	2	S78221		G-protein signalin
3	395.5	32.6	181	2	JC7228		G-protein signalin
4	383	31.5	519	2	S78089		G-protein signalin
5	345.5	28.5	196	2	S43436		B cell activation
6	293.5	24.2	251	2	S43576		C05B5.7 protein (c
7	293.5	24.2	254	2	G88571		protein C05B5.7 [i
8	291.5	24.0	181	2	T21035		hypothetical prote
9	291	24.0	169	2	T21034		hypothetical prote
10	290	23.9	244	2	T13580		hypothetical prote
11	277	22.8	173	2	S71812		RGS10 protein - hu
12	275.5	22.7	533	2	T31002		hypothetical prote
13	242	19.9	544	2	JC5503		G-protein signalin
14	234	19.3	1387	2	JC5502		G-protein signalin
15	217	17.9	558	2	T21468		hypothetical prote
16	216.5	17.8	473	2	T19337		hypothetical prote
17	193	15.9	838	2	T08423		Axin homolog Axil
18	192.5	15.9	832	2	T08422		negative regualtor
19	190.5	15.7	234	2	T26672		hypothetical prote
20	175	14.4	284	2	T15700		hypothetical prote
21	168.5	13.9	270	2	T22213		protein ZC53.7 [im
22	156	12.9	235	2	F89472		hypothetical prote
23	153	12.6	274	2	T22214		hypothetical prote
24	143	11.8	303	2	T29513		probable membrane
25	139.5	11.5	309	2	S61665		developmental regu
26	138	11.4	719	2	S60771		protein F52D2.2 [i
27	112.5	9.3	229	2	A89473		hypothetical prote
28	109.5	9.0	423	2	T15335		hypothetical prote
29	107.5	8.9	913	2	T31497		hypothetical prote
30	105.5	8.7	688	1	A39336		beta-adrenergic-re
31	104	8.6	1010	2	T41077		hypothetical struc
32	102	8.4	2748	2	S57976		nuclear migration
33	98.5	8.1	727	2	T26096		hypothetical prote
34	98	8.1	689	1	I56531		beta-adrenergic-re
35	98	8.0	2829	2	A42271		reticulocyte-bindi
36	97	8.0	688	1	JC1469		beta-adrenergic-re
37	96.5	7.9	901	2	S65161		hypothetical prote
38	93	7.7	704	2	S25371		DBF4 protein - yea
39	92	7.6	688	1	I73628		beta-adrenergic-re
40	92	7.6	792	2	T49989		hypothetical prote
41	92	7.6	1120	2	S67208		hypothetical prote
42	91.5	7.5	686	1	T15795		hypothetical prote
43	91.5	7.5	1031	1	A38713		kinesin heavy chai
44	91	7.5	449	2	G81419		hypothetical prote
45	91	7.5	851	2	S73939		oligopeptide trans

```
                              ALIGNMENTS

RESULT 1
I53020
G-0/G-1 switch regulatory protein 8 - human
N;Alternate names: helix-loop-helix phosphoprotein; regulator of G-protein signaling 2, ;
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C;Accession: I53020; I65984
R;Siderovski, D.P.; Heximer, S.P.; Forsdyke, D.R.
DNA Cell Biol. 13, 125-147, 1994
A;Title: A human gene encoding a putative basic helix-loop-helix phosphoprotein whose mRN
A;Reference number: I53020; MUID:94235158; PMID:8179820
A;Accession: I53020
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <RES>
A;Cross-references: GB:L13391; NID:g292036; PIDN:AAA20680.1; PID:g292037
A;Accession: I65984
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-211 <RE2>
A;Cross-references: GB:L13463; NID:g292054; PIDN:AAC37587.1; PID:g292055
C;Genetics:
A;Gene: GDB;RGS2; GOS8
A;Cross-references: GDB:355647; OMIM:600861
A;Map position: 1q31-1q31
A;Introns: 37/2; 71/2; 92/1; 147/3
C;Superfamily: B-cell activation protein BL34
C;Keywords: phosphoprotein

   Query Match             35.4%;  Score 429.5;  DB 2;  Length 211;
   Best Local Similarity   44.3%;  Pred. No. 4.6e-25;
   Matches   93;  Conservative   38;  Mismatches   70;  Indels    9;  Gaps    3;

Qy      4 SLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIR--AKEKRNRLSLLLQRPDFHGE  60
            : :|:|:|| : ||||    | ::| ::  :: ::|:|
Db      3 SAMFLAVQHDCRPMDKS-----AGSGHKSEBKREKMKRTLLKDWKTRLSYFLQNSSTPGK  57

Qy     61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
            | || |||| ::|| : :||::|| ||| |::|||| |:|:|||| ||||||||||||
Db     58 PKTGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACE 117

Qy    121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
            |||| |||| ||||:|| ||||||:  |||::|:|| |||  :: :  ||| |: :|:|
Db    118 DFKKTKSPQKLSSKARKLYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRV 177

Qy    181 YQLMEHDSYKRFLKSETVLHLIEGRPQRPT 210
            |||| | |||||||::  | :||| || |
Db    178 YSLMENNSYPRFLESEFYQDLCK-KPQITT 206


RESULT 2
S78221
```

G-protein signaling regulator RGP4 - human
N;Alternate names: G-protein signaling regulator RGS4
C;Species: Homo sapiens (man)
C;Date: 29-Jan-1998 #sequence_revision 13-Feb-1998 #text_change 31-Jan-2000
C;Accession: S78221; S68509
R;Druey, K.
submitted to the EMBL Data Library, May 1995
A;Reference number: S78089
A;Accession: S78221
A;Molecule type: mRNA
A;Residues: 1-205 <DRU>
A;Cross-references: EMBL:U27768; NID:g1216372; PIDN:AAC50395.1; PID:g1216373
R;Druey, K.M.; Blumer, K.J.; Kang, V.H.; Kehrl, J.H.
Nature 379, 742-746, 1996
A;Title: Inhibition of G-protein-mediated MAP kinase activation by a new mammalian gene
A;Reference number: A58012; MUID:96178495; PMID:8602223
A;Accession: S68509
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-35,'S',37-39,'S',41-46,'T',48-67,'T',69-76,'L',78-177,'T',179-197,'T',199
A;Cross-references: EMBL:U27768
A;Experimental source: brain
C;Superfamily: B-cell activation protein BL34

Query Match   32.7%;   Score 397.5;   DB 2;   Length 205;
Best Local Similarity   47.0%;   Pred. No. 1.e-22;
Matches   77;   Conservative   35;   Mismatches   49;   Indels   3;   Gaps   2;

```
Qy     41 AKEKRNRLSLLLQRPDF--HGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVD  98
             ||| ||:||:| ||:    || :| |:: ||:|| :||||||||||: :|| ||:
Db     16 AKDMKHRLGFLLQKSDSCEHNSSHNKKDKVVICQ-RVSQEBVVKWAESLENLISHECGLA  74

Qy     99 AFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTK 158
             || |||| |||||||||||||||| || :|||||||||||||||| | :|| |: |
Db     75 AFKAFLKSEYSEENIDFWISCEEYKKIKSPSKLSPKAKKIYNEFISVQATKEVNLDSCTR 134

Qy    159 EVIAKSIAQPTLHSFDTAQSRVQLMEHDSYKRFLKSETYLHLI 202
             |: | :|| ||||||||||||||||||||||||||||| |||:
Db    135 EETSRNMLEPTITCFDEAQKKIFNLMEKDSYRRFLKSRFYLDIV 178
```

RESULT 3
JC7228
G-protein signaling regulator 5 homolog - clawed frog
C;Species: Xenopus sp. (clawed frog)
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C;Accession: JC7228; PC7066
R;Saitoh, O.; Odagiri, M.; Masuho, I.; Nomoto, S.; Kinoshita, N.
Biochem. Biophys. Res. Commun. 270, 34-39, 2000
A;Title: Molecular cloning and characterization of Xenopus RGS5.
A;Reference number: JC7228
A;Accession: JC7228
A;Molecule type: mRNA
A;Residues: 1-181 <SAI>
A;Cross-references: DDBJ:AB018436
A;Accession: PC7066
A;Molecule type: protein
A;Residues: 88-168 <SA2>
A;Experimental source: embryo
C;Comment: This protein has function as a GTPase-activating protein. It interacts with a
genesis and cell differentiation.
C;Superfamily: B-cell activation protein BL34
C;Keywords: differentiation; embryo

Query Match   32.6%;   Score 395.5;   DB 2;   Length 181;
Best Local Similarity   46.0%;   Pred. No. 1.e-22;
Matches   81;   Conservative   31;   Mismatches   43;   Indels   21;   Gaps   2;

```
Qy     40 RAKEKRNRLSLLLQRPDFHGETQASRSAL---------LAKETRVSPEEAVKWAESF  87
             ||| ||:||:|:|
Db     15 RAKEIKTKLGTLLQKPE--------SAIDLIIPYPDKPEKPPKASKPSAEEAGQWRDSL  65

Qy     88 DKLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKKCKKEPQQIILKAKAIYEKFIQNDA 147
```

```
Db      66 EKLLENSYGLSVFQSFLKSEFSEENIEFWMACEDYKKAKSPSKMTTKAKKIYEEFIQTEA 125
            :||||||   |    ||  ||||||||| ||||:||||  | : :| |||||| ||| ||
Qy     148 PKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVQLMEHDSYKRFLKSETYLHLIE 203
            || :|   ::: |||||||||||||||||||||||||||||||||||||||||||
Db     126 PREVNIDHFTKAVTMKNLVEPSSASFELAQKKIFALMEKDSLPRFVRSEFYQELIK 181
```

RESULT 4
S78089
G-protein signaling regulator RGS3 - human
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 24-Nov-1999
C;Accession: S78089; S68436
R;Druey, K.
submitted to the EMBL Data Library, May 1995
A;Reference number: S78089
A;Accession: S78089
A;Molecule type: mRNA
A;Residues: 1-519 <DRU>
A;Cross-references: EMBL:U27655; NID:g1216368; PID:g1216369
R;Druey, K.M.; Blumer, K.J.; Kang, V.H.; Kehrl, J.H.
Nature 379, 742-746, 1996
A;Title: Inhibition of G-protein-mediated MAP kinase activation by a new mammalian gene f
A;Reference number: A58012; MUID:96178495; PMID:8602223
A;Accession: S68436
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-355,'K',357-519 <DRW>
A;Cross-references: EMBL:U27655
A;Experimental source: tonsil
A;Note: the sequence from Fig. 1b is inconsistent with that from Fig. 1a in having 356-Le

Query Match   31.5%;   Score 383.;   DB 2;   Length 519;
Best Local Similarity   43.7%;   Pred. No. 4e-21;
Matches   76;   Conservative   35;   Mismatches   63;   Indels   0;   Gaps   0;

```
Qy     36 EAKIRAKEKRNRLSLLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRD  95
Db    344 KSKNLAKDMKNKLGIFRRRNESPGAPPAGKADKMMSFKPFSEEAALKWGESLEKLLVHKY 403

Qy     96 GVDAFTRFLKTEFSEENIEFWVACEDFKKCKKEPQQIILKAKAIYEKFIQNDAPKEVNIDF 155
Db    404 GLAVFQAFLRTEFSEENLEFWLACEDFKKVKSQSKMASKAKKIFAEYIAIQACKEVNLDS 463

Qy    156 HTKEVIAKSIAQPTLHSFDTAQSRVQLMEHDSYKRFLKSETYLHLIEGRPQRP 209
Db    464 YTREHTKDNLQSVTRGCFDLAQKRIFGLMEKDSYPFLRSDLYLDLINQKKMSP 517
```

RESULT 5
S43436
B cell activation protein BL34 - human
N;Alternate names: B-cell activation protein 1r20; regulator of G-protein signaling 1 (RG
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Jan-2000
C;Accession: S43436; I56165; S34157
R;Newton, J.S.; Deed, R.W.; Mitchell, E.L.D.; Murphy, J.J.; Norton, J.D.
Biochim. Biophys. Acta 1216, 314-316, 1993
A;Title: A B cell specific immediate early human gene is located on chromosome band 1q31
A;Reference number: S43436; MUID:94060109; PMID:8241276
A;Accession: S43436
A;Molecule type: mRNA
A;Residues: 1-196 <NEW>
A;Cross-references: EMBL:X73427; NID:g313214; PIDN:CAA51826.1; PID:g313215
R;Hong, J.X.; Wilson, G.L.; Fox, C.H.; Kehrl, J.H.
J. Immunol. 150, 3895-3904, 1993
A;Title: Isolation and characterization of a novel B cell activation gene.
A;Reference number: I56165; MUID:93232596; PMID:8473738
A;Accession: I56165
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-180,'D',182-196 <HON>

A;Cross-references: GB:S59049; NID:g299704; PIDN:AAB26289.1; PID:G299705
C;Genetics:
A;Gene: GDB:RGS1; IER1; 1R20; IR20; BL34
A;Cross-references: GDB:439178; OMIM:600323
A;Map position: 1q31-1q31
C;Superfamily: B-cell activation protein BL34
C;keywords: B-cell; phosphoprotein

Query Match   28.5%; Score 345.5; DB 2; Length 196;
Best Local Similarity   52.3%; Pred. No. 7.9e-19;
Matches   67; Conservative   24; Mismatches   36; Indels   1; Gaps   1;

Qy   75 VSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILK 134
           |::|:| |||:|||||||||||||||||||:||| ||||||||||||||||||| |
Db   61 LSAAEVMQWSQSLEKLLANQTGQNVFGSFLKSEFSEENIEFWLACEDYKK-TESDLLPCK 119

Qy   135 AKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLK 194
           :||||||:: ||||||||||||||||||||||||||| ||| |||||
Db   120 AEEIYKAFVHSDAAKQINIDFRTRESTAKKIKAPTPTCFDEAQKVIYTLMEKDSYPRFLK 179

Qy   195 SETYLHLI 202
           ||| |:|:
Db   180 SHIYINLL 187

RESULT 6
S43576
C05B5.7 protein (clone C05B5) - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 13-Sep-1998
C;Accession: S43576
R;Mortimore, B.
submitted to the EMBL Data Library, April 1994
A;Reference number: S43570
A;Accession: S43576
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-251 <MOR>
C;Genetics:
A;Cross-references: EMBL:Z32679
A;Introns: 165/1; 197/3
C;Superfamily: B-cell activation protein BL34

Query Match   24.2%; Score 293.5; DB 2; Length 251;
Best Local Similarity   42.2%; Pred. No. 8e-15;
Matches   57; Conservative   30; Mismatches   37; Indels   11; Gaps   1;

Qy   75 VSPEEAVK--------WAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFK 123
           |::|:| |||:|||||||||||||||||||:||| ||||||||||||
Db   16 VSPQRSVQPEALSYEMVYSWQQSPDTLMSFKSGQKCFAEFLKSEYSDENILFWQACEELK 75

Qy   124 KCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVYQL 183
           ||| |||||| :|||||||:: ||||||||||||||||||||||||||| :||: |||||
Db   76 REKNSEKMEEKARIIYEDFISILSPKEVSLDSKVREIVNTNMSRPTQNTFEDAQHQIYQL 135

Qy   184 MEHDSYKRFLKSETY 198
           | | |||||||| |:
Db   136 MARDSYPRFLTSIFY 150

RESULT 7
G88571
protein C05B5.7 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-Aug-2001
C;Accession: G88571
R;anonymous,
Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: G88571

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-254 <STO>
A;Cross-references: GB:chr_III; PIDN:CAA83595.1; PID:g3873995; GSFPDB:GN00021; CESP:C05B5
A;Note: similar to 1r20 protein (human)
C;Genetics:
A;Gene: C05B5.7
A;Map position: 3
C;Superfamily: B-cell activation protein BL34

Query Match   24.2%; Score 293.5; DB 2; Length 254;
Best Local Similarity   42.2%; Pred. No. 8.1e-15;
Matches   57; Conservative   30; Mismatches   37; Indels   11; Gaps   1;

Qy   75 VSPEEAVK--------WAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFK 123
           |::|:| |||:|||||||||||||||||||:||| ||||||||||||
Db   19 VSPQRSVQPEALSYEMVYSWQQSPDTLMSFKSGQKCFAEFLKSEYSDENILFWQACEELK 78

Qy   124 KCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVYQL 183
           ||| |||||| :|||||||:: ||||||||||||||||||||||||||| :||: |||||
Db   79 REKNSEKMEEKARIIYEDFISILSPKEVSLDSKVREIVNTNMSRPTQNTFEDAQHQIYQL 138

Qy   184 MEHDSYKRFLKSETY 198
           | | |||||||| |:
Db   139 MARDSYPRFLTSIFY 153

RESULT 8
T21035
hypothetical protein F16H9.1b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C;Accession: T21035; T21272
R;Gardner, A.
submitted to the EMBL Data Library, July 1995
A;Reference number: Z19363
A;Accession: T21035
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-181 <WIL>
A;Cross-references: EMBL:Z50005; PIDN:CAB54219.1; GSPDB:GN00028; CESP:F16H9.1b
A;Experimental source: clone F16H9
R;Gardner, A.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19398
A;Accession: T21272
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-181 <WI2>
A;Cross-references: EMBL:Z67882; PIDN:CAB54230.1; GSPDB:GN00028; CESP:F16H9.1b
A;Experimental source: clone F22E10
C;Genetics:
A;Gene: CESP:F16H9.1b
A;Map position: X
A;Introns: 28/2; 43/2; 64/1; 103/3
C;Superfamily: B-cell activation protein BL34

Query Match   24.0%; Score 291.5; DB 2; Length 181;
Best Local Similarity   40.0%; Pred. No. 7.6e-15;
Matches   58; Conservative   31; Mismatches   51; Indels   5; Gaps   1;

Qy   59 GETQASRSALLAKETRVSP-----BEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENI 113
           | | || ||||||| || |||||||||||||||||||||||||||||||||
Db   23 GKPYVSGSVSVEKKNQENDGPPTYEIVFGWSQSFENLMKHRAGQKYFAEFLKGEYSDENI 82

Qy   114 EFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSF 173
           ||| ||||||||||| |||||||||||||:: |||||||||||||||||||||||||||
Db   83 LFWQACEELKREKNVAEKIEEKARIIYEDFISILSPKEVSLDSRVREIVTNMGRPSASTF 142

Qy   174 DTAQSRVYQLMEHDSYKRFLKSETY 198
           | || |||||| | ||||||||||
Db   143 DEAQNQIYTLMQRDSYPRFLIASNIY 167

RESULT 9
T21034
hypothetical protein F16H9.1a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C;Accession: T21034, T21270
R;Gardner, A.
submitted to the EMBL Data Library, July 1995
A;Accession: T21034
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-169 <WIL>
A;Cross-references: EMBL:Z50005; PIDN:CAA90295.1; GSPDB:GN00028; CESP:F16H9
A;Experimental source: clone F16H9
R;Gardner, A.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19398
A;Accession: T21270
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-169 <WI2>
A;Cross-references: EMBL:Z67882; PIDN:CAA91803.1; GSPDB:GN00028; CESP:F16H9.1a
A;Experimental source: clone F22E10
C;Genetics:
A;Gene: CESP:F16H9.1a
A;Map position: X
A;Introns: 16/2; 31/2; 52/1; 91/3
C;Superfamily: B-cell activation protein BL34

Query Match   24.0%;   Score 291; DB 2; Length 169;
Best Local Similarity 45.7%; Pred. No. 7.7e-15;
Matches   53; Conservative   26; Mismatches   37; Indels   0; Gaps   0;

Qy   83 WAESFDKLLSHRDGVDAFTRFLKTEFSSENIEFWVACEDFKKCKEPQQIILKAKAIYEKF 142
             :::|::|::| |::::| |||||||:|||||||:|| |::: |::|||:
Db   40 WSQSFENLMKGHRAGQKYFAEFLKGEYSDENILFWQACEELKREKNAEKIEEKARIIYEDF 99

Qy 143 IQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETY 198
             :::|::|: ||| :|::|||::||:|: |::|| ||||||:||:::||:|:
Db 100 ISILSPKEVSLDSRVREIVNTNMGRPSASTFDEAQNQIYTLMQRDSYPRFLASNIY 155

RESULT 10
T13580
hypothetical protein 52C10.2 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C;Accession: T13580
R;Benos, P.
submitted to the EMBL Data Library, February 1999
A;Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A;Reference number: Z17690
A;Accession: T13580
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-244 <BEN>
A;Cross-references: EMBL:AL035311; PIDN:CAA22951.1
C;Genetics:
A;Cross-references: FlyBase:FBgn0026309
A;Map position: X
A;Introns: 101/2; 127/2; 148/1
A;Note: EG:52C10.2
C;Superfamily: B-cell activation protein BL34

Query Match   23.9%;   Score 290; DB 2; Length 244;
Best Local Similarity 46.5%; Pred. No. 1.4e-14;
Matches   53; Conservative   23; Mismatches   38; Indels   0; Gaps   0;

Qy   78 EEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSSEENIEFWVACEDFKKCKEPQQIILKAKA 137
             :: :|:|:: |:||||||||||| ||||||||||:
Db 131 EEIRSWGKSFDKLMKSTAGRKVFQNFLRSEFSSENILFWLACEDLKKENSPELVEEKARL 190

Qy 138 IYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKR 191
             ||:|| |::|::|::||| ::||::|||::|||| |:: ||||||:||::||
Db 191 IYEDYISILSPREVSLDSRVREIVNRNMIEPTHTFDEAQIQIYTLMHRDSYPR 244

RESULT 11
S71812
RGS10 protein - human
C;Species: Homo sapiens (man)
C;Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 17-Mar-1999
C;Accession: S71812
R;Hunt, T.W.; Fields, T.A.; Casey, P.J.; Peralta, E.G.
Nature 383, 175-177, 1996
A;Title: RGS10 is a selective activator of Galpha(i) GTPase activity.
A;Reference number: S71812; MUID:96371048; PMID:8774883
A;Accession: S71812
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tran
A;Molecule type: mRNA
A;Residues: 1-173 <HUN>
C;Superfamily: B-cell activation protein BL34

Query Match   22.8%;   Score 277; DB 2; Length 173;
Best Local Similarity 36.5%; Pred. No. 8.7e-14;
Matches   62; Conservative   30; Mismatches   62; Indels   16; Gaps   3;

Qy   56 DFHGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSSENIEF 115
             | | ::: | : ||:|||| |:|||| |::
Db   12 DSDGSSSSSHQSL------------KSTAKWAASLENLLEDPEGVKRFREFLKEFSENVLF 62

Qy 116 WVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDT 175
             |:|||||| :: ||| |:|:||| |:: :|:|::|||
Db   63 WLACEDFKKMQDKTQMQEKAKEIYMTFLSSKASSQVNVEGQSR-LNEKILEBPHPLMFQK 121

Qy 176 AQSRVYQLMEHDSYKRFLKSETYL------HLIEGRPQRPTNLRRRSRSF 219
             |::|||||:|||| :|::::: |: ||: :: | :: ::
Db 122 LQDQIFNLMKDYSYSRFLKSDLFLKHKRTEBEEEEDLPDAQTAAKRASRIY 171

RESULT 12
T31002
hypothetical protein F56B6.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 29-Oct-1999
C;Accession: T31002
R;Stellye, L.; Stellye, L.
submitted to the EMBL Data Library, September 1999
A;Description: The sequence of C. elegans cosmid F56B6.
A;Reference number: Z20957
A;Accession: T31002
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-533 <STE>
A;Cross-references: EMBL:U64599; PIDN:AAB04563.1
A;Experimental source: strain Bristol N2
C;Genetics:
A;Map position: X
A;Introns: 39/3; 235/2; 259/2; 322/3; 360/2; 405/1; 461/3; 503/2
A;Note: F56B6.2

Query Match   22.7%;   Score 275.5; DB 2; Length 533;
Best Local Similarity 28.7%; Pred. No. 4.3e-13;
Matches   70; Conservative   44; Mismatches   79; Indels   51; Gaps   6;

Qy   15 ESKEKTFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGETQASRS--------- 66
             :::::| :|:: | : : : |: | ::: :::
Db 299 EINNEGFFLLNDQGRKQNFPQRKVRRQ------TFYEDDAFTGVSSAGGSVISNNTGG 351

Qy   67 ------------------------------ALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFT 101

Db 352 QVRIALKVIGPISRTLSYLRSKMDLALSTSSLYPSRDDVRQWEISFESLLNNKFGCALFR 411

Qy 102 RFLKTEFSSEENIEFWVACEDFKKCKE-PQQIIKAKAIYEKFIQNDAPKEVNIDFHTKEV 160
             :|||||:::||| ::
Db 102 RFLKTEFSEENIEFWVACEDFKKCKE-PQQIIKAKAIYEKFIQNDAPKEVNIDFHTKEV 160

```
Matches   80;  Conservative   35;  Mismatches   99;  Indels   46;  Gaps   9;

Qy    3 MSLVFFSQLNMCESKE-KTFFKLMHGSGKEETSIEAKIRAKEKRN-RLSLLLQRPDFHGE  60
          :|: |||:|::::| |.:::|:| :::..||||:::| ::
Db  609 MQNLFGPHRNVRKTKEDKKSSKLGRGVALAQTSQRTSARRSFGRSRRFSLTRSLDDLESA 668

Qy   61 T----------------------QASRSALLAKET--RVSPEEAVKWAESFDKLLSHRDGVDA  99
        |                       ::|:  :::  :::|||  ::||||||||||
Db  669 TVSDGELTGADLKDCISNNSLSSNASLPSVQSCRRLRERRVASWAVSFERLLQDPVGVRY 728

Qy  100 FTRFLKTEFSEENIEFWVACEDFK--KCKEPQQILKAKAIYEKFIQNDAPKEVNIDFHT  157
        |: ||:|:|:||||:|||:||| |   ::|  :: |:::: :|| :| :::|||:
Db  729 FSDFLRKFKFSEENILFWQACECFSHVPAHDKKELSYRAREIFSKFLCSKATTPVNID--S 786

Qy  158 KEVIAKSIAQ-PTLHSFDTAQSRVYQLMEHDSYKRFLKSETY----LHLIEGRPQRPTNL 212
         :.:|: :: | ::|   ::|||:|||:|||||.||||| || :|::|:| :
Db  787 QAQLADDIINAPHPDMFKEQQLQIFNLMKFDSYTRFLKSQLYQECVLABVEGR-------- 839

Qy  213 RRRSRSFTYNDFQDVKSDVA 232
Db  840 --------TLPDSQQVPSSPA 852


RESULT 15
T21468
hypothetical protein F28C1.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21468
R;Kershaw, J.
submitted to the EMBL Data Library, June 1996
A;Reference number: 219427
A;Accession: T21468
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-558 <WIL>
A;Cross-references: EMBL:Z75539; PIDN:CAA99844.1; GSPDB:GN00023; CESP:F28C1.2
A;Experimental source: clone F28C1
C;Genetics:
C;Gene: CESP:F28C1.2
A;Map position: 5
A;Introns: 7/2; 76/1; 129/1; 176/2; 313/2; 364/2; 409/2; 515/3

Query Match           17.9%;  Score  217;  DB 2;  Length 558;
Best Local Similarity 33.1%;  Pred. No. 1.1e-08;
Matches   50;  Conservative  24;  Mismatches  61;  Indels  16;  Gaps  3;

Qy   53 QRPDFHGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDATRFLKTEFSEEN 112
Db  405 QHDKITGDIQTRLKL------------------WEDSFEBLLADSLGRETLQKFLDKEYSGEN 450

Qy  113 IEFWVACEDFKKCKEPQQILKAKAIYEKFIQ-NDAPKEVNIDFHTKEVIAKSIAQPTLH 171
Db  451 LRFWEVQXLRKCSS-RMVPVMVTEIYNEFIDTNAATSPVNVDCKVMEVTEDNLKNPNRW 509

Qy  172 SFDTAQSRVYQLMEHDSYKRFLKSETYLHLI 202
Db  510 SFDEAADHIYCLMKQNDSYQRFLRSEIYKDLV 540


Search completed: August 20, 2004, 16:47:49
Job time : 26 secs
```

```
Db  412 QFLKKEFSDENMDFWLECEEFKRMKDGKKSTTQKAIEIYSEFVAEHSPKEVNLDSDTRAA 471

Qy  161 IAKSI---AQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEG----RPQRPTNL 212
Db  472 TKAAVEAGCKP--DTFALAQSRVEQLMSKDSYRRFLRDRLFLDLLESYEITDKEDKPSSS 529

Qy  213 RRRS 216
Db  530 KDKN 533


RESULT 13
JC5503
G-protein signaling regulator 14 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 05-Nov-1999
C;Accession: JC5503
R;Snow, B.E.; Antonio, L.; Suggs, S.; Gutstein, H.B.; Siderovski, D.P.
Biochem. Biophys. Res. Commun. 233, 770-777, 1997
A;Title: Molecular cloning and expression analysis of rat Rgs12 and Rgs14.
A;Reference number: JC5502; MUID:97312490; PMID:9168931
A;Accession: JC5503
A;Molecule type: mRNA
A;Residues: 1-544 <SNO>
A;Cross-references: GB:U92279; NID:g2088555; PIDN:AAC53175.1; PID:g2088556
C;Comment: This protein functions as GTPase activating protein.
F;64-113/Domain: GH1 #status predicted <GH1>
F;117-152/Domain: GH2 #status predicted <GH2>
F;155-180/Domain: GH3 #status predicted <GH3>
F;514-534/Region: conserved #status predicted

Query Match           19.9%;  Score 242;  DB 2;  Length 544;
Best Local Similarity 36.1%;  Pred. No. 1.4e-10;
Matches   57;  Conservative  28;  Mismatches  51;  Indels  22;  Gaps  5;

Qy   78 EEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKK--CKEPQQIILKA 135
Db   59 QPVASWAQSFERLLQDPRGLAYFTEFLKKEFSAENVTFWQACERFQQIPASDTKQLAQEA 118

Qy  136 KAIYEKFIQNDAPKEVNID---FHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRF 192
Db  119 HNIYHEFLSSQALSPVNIDRQAWLSEEV----LAQPRPDMFRAQQLQIFNLMKFDSYARF 174

Qy  193 LKSETYLHLI----EGRPQR-------PTNLRRRSR 217
Db  175 VKSPLYQECVLABAEGRPLREPGSSHLGSPDTARKKPK 212


RESULT 14
JC5502
G-protein signaling regulator 12 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 05-Nov-1999
C;Accession: JC5502
R;Snow, B.E.; Antonio, L.; Suggs, S.; Gutstein, H.B.; Siderovski, D.P.
Biochem. Biophys. Res. Commun. 233, 770-777, 1997
A;Title: Molecular cloning and expression analysis of rat Rgs12 and Rgs14.
A;Reference number: JC5502; MUID:97312490; PMID:9168931
A;Accession: JC5502
A;Molecule type: mRNA
A;Residues: 1-1387 <SNO>
A;Cross-references: GB:U92280; NID:g2088557; PIDN:AAC53176.1; PID:g2088558
C;Comment: This protein functions as GTPase activating protein. It interacts with ras-li
F;18-80/Domain: rhophilin-like #status predicted <RHO>
F;712-761/Domain: GH1 #status predicted <GH1>
F;765-800/Domain: GH2 #status predicted <GH2>
F;804-828/Domain: GH3 #status predicted <GH3>
F;1204-1220/Region: conserved #status predicted
F;1266-1295/Region: coiled heptad repeat (S-P-X-S-A)

Query Match           19.3%;  Score 234;  DB 2;  Length 1387;
Best Local Similarity 30.8%;  Pred. No. 1.7e-09;
```

```
              Copyright    GenCore version 5.1.6
                           Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      August 20, 2004, 12:32:00 ; Search time 14.5 Seconds
                                          (without alignments)
                                          843.896 Million cell updates/sec

Title:       US-09-894-749-4
Perfect score: 1214
Sequence:    1 MDMSLVFFSQLNMCESKEKT........SRSFTYNDFQDVKSDVAIWL 235

Scoring table: BLOSUM62
             Gapop 10.0 , Gapext 0.5

Searched:    141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters:    141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                          SUMMARIES
```

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1214	100.0	235	1	RGSI_MOUSE	Q99pg4 mus musculu
2	1032	85.0	235	1	RGSI_HUMAN	Q9ns28 homo sapien
3	429.5	35.4	211	1	RGS2_HUMAN	P41220 homo sapien
4	428.5	35.3	211	1	RGS2_RAT	Q9jhx0 rattus norv
5	427.5	35.1	181	1	RGS2_HUMAN	O15539 homo sapien
6	426.5	35.1	211	1	RGS2_MOUSE	O08849 mus musculu
7	412.5	34.0	181	1	RGS5_RAT	P49800 rattus norv
8	407.5	33.6	180	1	RGS8_HUMAN	P57771 homo sapien
9	407.5	33.6	180	1	RGS8_RAT	P49804 rattus norv
10	406.5	33.5	181	1	RGS5_MOUSE	O08850 mus musculu
11	397.5	32.7	205	1	RGS4_HUMAN	P49798 homo sapien
12	397.5	32.7	205	1	RGS4_MOUSE	O08899 mus musculu
13	396	32.6	568	1	RGS3_MOUSE	Q9dc04 mus musculu
14	395.5	32.6	205	1	RGS4_RAT	P49799 rattus norv
15	383	31.5	519	1	RGS3_HUMAN	P49796 homo sapien
16	367.5	30.3	196	1	RGSI_MOUSE	Q9j125 mus musculu
17	363	29.9	202	1	RGSG_HUMAN	O15492 homo sapien
18	356.5	29.4	199	1	RGSG_RAT	P56700 rattus norv
19	356	29.3	201	1	RGSG_MOUSE	P97428 mus musculu
20	356	29.3	202	1	RGSG_BOVIN	O46471 bos taurus
21	347.5	28.6	196	1	RGSI_HUMAN	Q08116 homo sapien
22	311.5	25.7	216	1	RGSJ_RAT	O70521 rattus norv
23	309	25.5	374	1	RGSK_BOVIN	P79948 bos taurus
24	308	25.4	388	1	RGSK_HUMAN	O76081 homo sapien
25	307	25.3	217	1	RGSJ_HUMAN	P49795 homo sapien
26	306.5	25.2	216	1	RGSJ_MOUSE	Q9cx84 mus musculu
27	302	24.9	218	1	RGSK_CHICK	Q9pwal gallus gall
28	300	24.7	159	1	RGSD_HUMAN	O14921 homo sapien
29	300	24.7	239	1	RGSK_MOUSE	Q9qzb1 mus musculu
30	298	24.5	210	1	RGSH_CHICK	Q9pwa0 gallus gall
31	298	24.5	210	1	RGSH_HUMAN	Q9ugc6 homo sapien
32	295	24.3	210	1	RGSH_MOUSE	Q9qb0 mus musculu
33	293.5	24.2	201	1	RGS1_CAEEL	P34295 caenorhabdi
34	291	24.0	169	1	RGS2_CAEEL	P49808 caenorhabdi
35	286.5	23.6	181	1	RGSA_MOUSE	Q9cqe5 mus musculu
36	277	22.8	173	1	RGSA_HUMAN	O43665 homo sapien
37	259.5	21.4	409	1	RGS6_MOUSE	Q9z2h2 mus musculu
38	248	20.4	677	1	RGS9_RAT	P49805 rattus norv
39	247	20.3	675	1	RGS9_MOUSE	O54828 mus musculu
40	246	20.3	484	1	RGS9_BOVIN	O46469 bos taurus
41	246	20.3	674	1	RGS9_HUMAN	O75916 homo sapien
42	245.5	20.2	495	1	RGS7_HUMAN	P49802 homo sapien
43	243	20.0	547	1	RGSE_MOUSE	P97492 mus musculu
44	242	19.9	544	1	RGSE_RAT	O08773 rattus norv
45	241.5	19.9	567	1	RGS6_HUMAN	P49758 homo sapien

```
                          ALIGNMENTS

RESULT 1
RGSI_MOUSE
ID   RGSI_MOUSE    STANDARD;    PRT;    235 AA.
AC   Q99PG4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Regulator of G-protein signaling 18 (RGS18).
GN   RGS18.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCSI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BA;
RX   PubMed=1042171;
RA   Park I.K., Klug C.A., Li K., Jerabek L., Li L., Nanamori M.,
RA   Neubig R.R., Hood L., Weissman I.L., Clarke M.F.;
RT   "Molecular cloning and characterization of a novel regulator of
RT   G-protein signaling from mouse hematopoietic stem cells.";
RL   J. Biol. Chem. 276:915-923(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BDF1;
RX   MEDLINE=21240406; PubMed=11342430;
RA   Nagata Y., Oda M., Nakata H., Shozaki Y., Kozasa T., Todokoro K.;
RT   "A novel regulator of G-protein signaling bearing GAP activity for
RT   Galphai and Galphaq in megakaryocytes.";
RL   Blood 97:3051-3060(2001).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. BINDS TO G(I)-ALPHA AND G(Q)-ALPHA.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: Expressed in bone marrow, spleen, fetal liver
CC       and lung. At very low levels expressed in heart.
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   ------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------
DR   EMBL; AF302685; AAG53657.1; -.
DR   EMBL; AB042807; BAB60723.1; -.
DR   HSSP; P49799; 1AGR.
DR   MGD; MGI:1927498; Rgs18.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005096; F:GTPase activator activity; IDA.
DR   GO; GO:0016299; F:regulator of G-protein signaling activity; IDA.
DR   GO; GO:0008277; P:regulation of G-protein coupled receptor pr. ..; IDA.
DR   InterPro; IPR000342; RegI_Gprotein.
DR   Pfam; PF00615; RGS; 1.
```

```
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Reg1_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor. RGS.
FT   DOMAIN       86    202       RGS.
SQ   SEQUENCE   235 AA;  27610 MW;  643343SAC1CBCFD8 CRC64;

Query Match             100.0%;  Score 1214;  DB 1;  Length 235;
Best Local Similarity   100.0%;  Pred. No. 1.7e-84;
Matches 235;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDMSLVFFSQLNMCESKEKTFFKLMEGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60

Qy      61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWIACE 120

Qy     121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180

Qy     181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235

RESULT 2
RGS1_HUMAN
ID   RGS1_HUMAN     STANDARD;     PRT;    235 AA.
AC   Q9NS28;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Regulator of G-protein signaling 18 (RGS18).
GN   RGS18 OR RGS13.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhang W., Wan T., Yuan Z., He L., Cao X.;
RT   "A novel regulator of G-protein signaling.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Platelet;
RA   Gagnon A.W., Murray D.L., Leadley R.J. Jr.;
RT   "cloning and characterization of a novel regulator of G-protein
RT   signaling in human platelets.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   PubMed=11042171;
RA   Park I.K., Klug C.A., Li K., Jerabek L., Li L., Nanamori M.,
RA   Neubig R.R., Hood L., Weissman I.L., Clarke M.F.;
RT   "Molecular cloning and characterization of a novel regulator of
RT   G-protein signaling from mouse hematopoietic stem cells.";
RL   J. Biol. Chem. 276:915-923 (2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. BINDS TO G(I)-ALPHA AND G(Q)-ALPHA
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in peripheral leukocytes, bone
CC       marrow, spleen and fetal liver.
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF076642; AAF80227.1; -.
DR   EMBL; AF268036; AAK58589.1; -.
DR   EMBL; BC026632; AAH20632.1; -.
DR   HSSP; P49799; 1AGR.
DR   Genew; HGNC:14261; RGS18.
DR   MIM; 607192; -.
DR   InterPro; IPR000342; Reg1_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Reg1_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor.    RGS.
FT   DOMAIN       86    202       RGS.
FT   CONFLICT    226    227       DV -> ML (IN REF. 3).
SQ   SEQUENCE   235 AA;  27582 MW;  973ABD8BC7DE3D5 CRC64;

Query Match             85.0%;  Score 1032;  DB 1;  Length 235;
Best Local Similarity   83.8%;  Pred. No. 8.2e-71;
Matches 197;  Conservative  19;  Mismatches  19;  Indels  0;  Gaps  0;

Qy       1 MDMSLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIRAKEKRNRLSLLLQRPDFHGE 60
           ||||||||: |:|||||||||||||: ||||||||:||||||||||||||||:| |: |
Db       1 METTLLFFSQINMCESKEKTFFKLIHGSGKEETSKEAKIRAKEKRNRLSLLVQKPEFHED 60

Qy      61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
           |::||| ||:|||||||||||||| ::||||||||||: ||||||||||||||||||:||
Db      61 TRSSRSGHLAKETRVSPEEAVKWGSSFDKLLSHRDGLEAFTRFLKTEFSEENIEFWIACE 120

Qy     121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
           ||||  | |||:|||||||||||| |||||||:|||||||::|:| |||||||||:|||||
Db     121 DFKKKSKGPQQIHLKAKAIYEKFIQTDAPKEVNLDFHTKEVTNSITQPTLHSFDAAQSRV 180

Qy     181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFTYNDFQDVKSDVAIWL 235
           |||||:||| ||||| :||| |:|||||||||||||||||||::||||:|||||||
Db     181 YQLMEQDSYTRFLKSDIYLDLMEGRPQRPTNLRRRSRSF*CNEFQDVQSDVAIWL 235

RESULT 3
RGS2_HUMAN
ID   RGS2_HUMAN     STANDARD;     PRT;    211 AA.
AC   P41220;
DT   01-FEB-1995 (Rel. 31, Created)
```

```
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Regulator of G-protein signaling 2 (RGS2) (G0/G1 switch regulatory
DE   protein 8).
GN   RGS2 OR GOS8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=94235158; PubMed=8179820;
RA   Siderovski D.P., Heximer S.P., Forsdyke D.R.;
RT   "A human gene encoding a putative basic helix-loop-helix
RT   phosphoprotein whose mRNA increases rapidly in cycloheximide-treated
RT   blood mononuclear cells.";
RL   DNA Cell Biol. 13:125-147(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
RT   "cDNA clones of human proteins involved in signal transduction
RT   sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Bagguley C.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Urinary bladder;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H.F., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   CHARACTERIZATION.
RX   MEDLINE=95371353; PubMed=7643615;
RA   Wu H.-K., Heng H.H., Shi X.-M., Forsdyke D.R., Tsui L.-C., Mak T.W.,
RA   Minden M.D., Siderovski D.P.;
RT   "Differential expression of a basic helix-loop-helix phosphoprotein
RT   gene, GOS8, in acute leukemia and localization to human chromosome
RT   1q31.";
RL   Leukemia 9:1291-1298(1995).
RN   [6]
RP   PHOSPHORYLATION.
RX   MEDLINE=21101914; PubMed=11063746;
RA   Cunningham M.L., Waldo G.L., Hollinger S., Hepler J.R., Harden T.K.;
RT   "Protein kinase C phosphorylates RGS2 and modulates its capacity for
RT   negative regulation of Galpha 11 signaling.";
RL   J. Biol. Chem. 276:5438-5444(2001).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. MAY PLAY A ROLE IN LEUKEMOGENESIS.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ACUTE MYELOGENOUS LEUKEMIA (AML)
CC       AND IN ACUTE LYMPHOBLASTIC LEUKEMIA (ALL).
```

```
CC   -!- PTM: Phosphorylated by protein kinase C.
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L13391; AAA20680.1; -.
DR   EMBL; L13463; AAC37587.1; -.
DR   EMBL; AF493926; AAM12640.1; -.
DR   EMBL; AL035407; CAB62512.1; -.
DR   EMBL; BC007049; AAH07049.1; -.
DR   PIR; I53020; I53020.
DR   HSSP; P49799; 1AGR.
DR   Genew; HGNC:9998; RGS2.
DR   MIM; 600861; -.
DR   GO; GO:0005516; F:calmodulin binding; TAS.
DR   GO; GO:0005096; F:GTPase activator activity; TAS.
DR   GO; GO:0008277; P:regulation of G-protein coupled receptor pr...; TAS.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Cell cycle; Phosphorylation.
FT   DOMAIN       83    199       RGS.
SQ   SEQUENCE   211 AA;  24382 MW;  EFFE4AE47EF9AD8F CRC64;

Query Match             35.4%;  Score 429.5;  DB 1;  Length 211;
Best Local Similarity   44.3%;  Pred. No. 1.4e-25;
Matches   93;  Conservative   38;  Mismatches  70;  Indels   9;  Gaps   3;

Qy      4 SLVFFSQLMCESKEKTFFKLMHGSGKETSIEAKIR--AKEKRNRLSLLLQRPDFHGE   60
            |  || | :| :| |   | |::|  ||   |     |  |  |  |   | |||
Db      3 SAMFLAVQHDCRPMDKS-----AGSGHKSEERKEMKRTLLKDWKTRLSYFLQNSSTPGK   57

Qy     61 TQASRALLAKETRVSPEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE  120
            | |: | |||  :  | |   ||: |  | | |  || | |  ||| :| | |||| |
Db     58 PKTGKKSKQQAFIKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACE  117

Qy    121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV  180
            |||| | || |  ||||| ||  :| :|  |  |  | |  | ||  |  |  |  | ||
Db    118 DFKKTKSPQKLSSKARKIYTDFIEKEAPEINIDFQTKTLIAQNIQEATSGCFTTAQXRV  177

Qy    181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPT  210
            | |  ||| ||||:|| | | ||||| |
Db    178 YSLMENNSYPRFLESEFYQDLCK-KPQITT  206
```

RESULT 4
```
RGS2_RAT
ID   RGS2_RAT      STANDARD;      PRT;    211 AA.
AC   Q9DHX0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Regulator of G-protein signaling 2 (RGS2).
GN   RGS2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Thoracic aorta;
RX   MEDLINE=20150060; PubMed=10692485;
RA   Grant S.L., Lassegue B., Griendling K.K., Ushio-Fukai M., Lyons P.R.,
RA   Alexander R.W.;
```

```
RT   "Specific regulation of RGS2 messenger RNA by angiotensin II in
RT   cultured vascular smooth muscle cells.";
RL   Mol. Pharmacol. 57:460-467(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Brain;
RA   Taymans J.-M., Wintmolders C., Te Riele P., Jurzak M.,
RA   Groenewegen H.J., Leysen J.E., Langlois X.;
RT   "A detailed distribution study of RGS2 messenger RNA and protein in
RT   the rat brain.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Ko J.K., Kim I.S., Park D.H.;
RT   "PTH induced RGS cDNA sequence in rat osteoblast-like UMR106 cell.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 1-100 FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=21302414; PubMed=11409749;
RA   Robinet E.A., Wurch T., Pauwels P.J.;
RT   "Haloperidol and clozapine differently affect RGS2 mRNA expression in
RT   rat brain.";
RL   NeuroReport 12:1731-1735(2001).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. MAY PLAY A ROLE IN LEUKEMOGENESIS
CC       (BY SIMILARITY).
CC   -!- PTM: Phosphorylated by protein kinase C (By similarity).
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF279918; AAF85981.1; -.
DR   EMBL; AF321837; AAK09375.1; -.
DR   EMBL; AY043246; AAK85309.1; -.
DR   EMBL; AJ318489; CAC44900.1; -.
DR   HSSP; P49799; 1AGR.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Cell cycle; Phosphorylation.
FT   DOMAIN       83    199       RGS.
SQ   SEQUENCE   211 AA;  24323 MW;  84B5525CB41EA964 CRC64;
```

```
Query Match             35.3%;    Score 428.5;    DB 1;    Length 211;
Best Local Similarity   44.3%;    Pred. No. 1.7e-25;
Matches  93;  Conservative  39;  Mismatches  69;  Indels  9;  Gaps  3;

Qy      4  SLVFFSQLNMCESKEKTFFKLMHGSGKEETSTEAKIR--AKEKRNRLSLLLQRPDFHGE   60
           : |:: ||: ||::|  | :||||::  : |:   |::   ::| ||:
Db      3  SAMFLAVQHDCVPMDKS----AGNGPKVEEKREKMKRTLLKDWKTRLSYFLQNSSTPGK   57

Qy     61  TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRPLKTEPSEENIEFWVACE  120
           |  :: ||:| |  |||||| |||: ||:||:|:|| | |::|||: ||||| |||:|||
Db     58  PKTGKKSKQQTFIKPSPEALLWAEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACE  117

Qy    121  DFKKCKEPQOIILKAKAIYEKFIQNDAPKEVNIDFHKEVIAKSIAQPTLHSFDTAQSRV  180
           |||||:|  ||:|:||  :||:|::| || |::||:||| :||  |||:| ||:||||:
Db    118  DFKKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKTLIAQNIQEATSGCFTTAQKRV  177

Qy    181  YQLMEHDSYKRFLKSETYLHLIEGRPQRPT  210
           |||||::| ||| |||| ||::|:|| |:|
Db    178  YSLMENNSYPRFLESEFYQDLCK-KPQITT  206
```

```
RESULT 5
RGS5_HUMAN
ID   RGS5_HUMAN    STANDARD;    PRT;    181 AA.
AC   O15539;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Regulator of G-protein signaling 5 (RGS5).
GN   RGS5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Chatterjee T.K., Fisher R.A.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Neuroblastoma;
RX   MEDLINE=98419174; PubMed=9747037;
RA   Seki N., Sugano S., Suzuki Y.-I., Nakagawara A., Ohira M.,
RA   Muramatsu M.-A., Saito T., Hori T.;
RT   "Isolation, tissue expression, and chromosomal assignment of human
RT   RGS5, a novel G-protein signaling regulator gene.";
RL   J. Hum. Genet. 43:202-205(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=99403338; PubMed=10471929;
RA   Cismowski M.J., Takesono A., Ma C., Lizano J.S., Xie X.,
RA   Fuernkranz H., Lanier S.M., Duzic E.;
RT   "Genetic screens in yeast to identify mammalian nonreceptor modulators
RT   of G-protein signaling.";
RL   Nat. Biotechnol. 17:878-883(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
RT   "cDNA clones of human proteins involved in signal transduction
RT   sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McSwan J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. BINDS TO G(I)-ALPHA AND G(O)-ALPHA,
CC       BUT NOT TO G(S)-ALPHA (BY SIMILARITY).
CC   -!- SIMILARITY: Contains 1 RGS domain.
```

CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AF030108; AAB84001.1; -.
DR   EMBL; AB008109; BAA22889.1; -.
DR   EMBL; AF159570; AAD40957.1; -.
DR   EMBL; AF493929; AAM12643.1; -.
DR   EMBL; BC030059; AAH30059.1; -.
DR   HSSP; P49799; 1AGR.
DR   Genew; HGNC:10001; RGS5.
DR   MIM; 603276; -.
DR   GO; GO:0005096; F:GTPase activator activity; TAS.
DR   GO; GO:0008277; P:regulation of G-protein coupled receptor pr... .; TAS.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor.
FT   DOMAIN   64   180   RGS.
SQ   SEQUENCE   181 AA; 20946 MW; 2E08CB0179DE7687 CRC64;

Query Match   35.2%; Score 427.5; DB 1; Length 181;
Best Local Similarity 50.3%; Pred. No. 1.7e-25;
Matches   84; Conservative   33; Mismatches   47; Indels   3; Gaps   1;

Qy   40 RAKEKRNRLSLLLQRPDFHGE--TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDG 96
          |||| | || | || | | || : | : ::||: ::|| || ||:|||:|
Db   15 RAKEIKIKLGILLQKPDSVGDLVIPYNEKPERPAKTQKTSLDEALQWRDSLDKLLQNNYG 74

Qy   97 VDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFH 156
          | || |||| :||| | : ::| |||||||||||||||||||||||||||||||||
Db   75 LASFKSFLKSEFSENLEFWIACEDYKKIKSPAKWAEKAKQIYEEFIQTEAPKEVNIDHF 134

Qy   157 TKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIE 203
          ||: ||||: ::||| |||||||||||||:| |||||||| |
Db   135 TKDITMKNLVEPSLSSFDMAQKRIHALMEKDSLPRFVRSEFYQELIK 181

RESULT 6
RGS2_MOUSE
ID   RGS2_MOUSE   STANDARD;   PRT;   211 AA.
AC   CO0849; Q91WX1; Q9JL24;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Regulator of G-protein signaling 2 (RGS2).
GN   RGS2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97236828; PubMed=9079700;
RA   Chen C., Zheng B., Han J., Lin S.C.;
RT   "Characterization of a novel mammalian RGS protein that binds to
RT   Galpha proteins and inhibits pheromone signaling in yeast.";
RL   J. Biol. Chem. 272:8679-8685(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=20243574; PubMed=10779778;
RA   Reif K., Cyster J.G.;
RT   "RGS molecule expression in murine B lymphocytes and ability to
RT   down-regulate chemotaxis to lymphoid chemokines.";
RL   J. Immunol. 164:4720-4429(2000).

RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RA   Shen Q.-X., Wang J., Huang Z.-P.;
RT   "Identification of novel endometrial and embryonic factors involved in
RT   mouse embryo implantation.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC   ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC   THEIR INACTIVE GDP-BOUND FORM. MAY PLAY A ROLE IN LEUKEMOGENESIS.
CC   -!- TISSUE SPECIFICITY: Expressed in a wide variety of tissues.
CC   -!- PTM: Phosphorylated by protein kinase C (By similarity).
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; U67187; AAB50617.1; -.
DR   EMBL; AF215668; AAF34625.1; -.
DR   EMBL; AF432916; AAL28114.1; -.
DR   HSSP; P49799; 1AGR.
DR   MGD; MGI:1098271; Rgs2.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Cell cycle; Phosphorylation.
FT   DOMAIN   83   199   RGS.
FT   CONFLICT   39   40   KD -> NH (IN REF. 1).
FT   CONFLICT   78   79   QL -> HV (IN REF. 1).
FT   CONFLICT   78   79   QL -> DV (IN REF. 3).
SQ   SEQUENCE   211 AA; 24294 MW; 5D63255C2BC7E7FA CRC64;

Query Match   35.1%; Score 426.5; DB 1; Length 211;
Best Local Similarity 44.3%; Pred. No. 2.4e-25;
Matches   93; Conservative   38; Mismatches   70; Indels   9; Gaps   3;

Qy   4 SLVFFSQLNMCESKEKTFFKLMHGSGKEETSIEAKIR--AKEKRNRLSLLLQRPDFHGE 60
          : || | | :: : ||: : :|| : | : | :: || | || |
Db   3 SAMFLAVDHDCVPMDKS----AGNGPKVEEKREKMRTLLKDWKTRLSYFLQNSSAPGK 57

Qy   61 TQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACE 120
          ||: | : ::| || | || ||:||| |||| ||| :||| | :
Db   58 PKTGKKSKQQTFIKPSPEEAQLWAEAFDELLASKYGLAAFRAFLKSEFCEENIEFWLACE 117

Qy   121 DFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRV 180
          |||| ||| |:::| ||::| ||| ||| ||||| : : ||: ||
Db   118 DFKKTKSPQKLSSKARKIYTDFIEKEAPKEINIDFQTKSLIAQNIQEATSGCFTTAQKRV 177

Qy   181 YQLMEHDSYKRFLKSETYLHLIEGRPQRPT 210
          ||| |||:| ||:|| | | :
Db   178 YSIMENNSYPRFLESEFYQDLCK-KPQITT 206

RESULT 7
RGS5_RAT
ID   RGS5_RAT   STANDARD;   PRT;   181 AA.
AC   P49800; Q9JKD7;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Regulator of G-protein signaling 5 (RGS5).
GN   RGS5.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kirsch T., Lippoldt A., Wellner M., Haller H.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 95-161 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=96146645; PubMed=8548815;
RA   Koelle M.R., Horvitz H.R.;
RT   "EGL-10 regulates G protein signaling in the C. elegans nervous
RT   system and shares a conserved domain with many mammalian proteins.";
RL   Cell 84:115-125(1996).
-!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. BINDS TO G(I)-ALPHA AND G(O)-ALPHA,
CC       BUT NOT TO G(S)-ALPHA (BY SIMILARITY).
-!- SIMILARITY: Contains 1 RGS domain.

CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).

DR   EMBL; AF241259; AAF73424.1; -.
DR   EMBL; U32435; AAC52372.1; -.
DR   HSSP; P49799; 1AGR.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor.
FT   DOMAIN       64    180       RGS.
SQ   SEQUENCE   181 AA;  21052 MW;  9C179CE008B82C03 CRC64;

Query Match             34.0%;  Score 412.5;  DB 1;  Length 181;
Best Local Similarity   49.1%;  Pred. No. 2.2e-24;
Matches   86;  Conservative   27;  Mismatches   41;  Indels   21;  Gaps    2;

Qy    40  RAKEKRNRLSLLLQRPDFHGETQASRSAL------------LAKETRVSPEBAVKWAESF  87
          ||||||  ||||||||         |||               |||   |||||||||
Db    15  RAKEIKIKLGILLQKPD--------SAVDLVIPYNEKPEKPAKAHKPSLEEVLQWRQSL  65

Qy    88  DKLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDA  147
          ||||     |||||||||||| ||  ||||      ||||||||||||||  |
Db    66  DKLLQSNYGFASFKSFLKSEFSENLEFWVACENYKKIIKSPIIMAEKAKQIYEFIQTEA  125

Qy    148 PKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLI  202
          |||||| ||||        |||||||   ||||||||||||||| ||||
Db    126 PKEVNIDHFTKDITMKNLVEPSPHSFDLAQKRIYALMEKDSLPRFVRSEFYKELI  180


RESULT 8
ID   RGS8_HUMAN              STANDARD;      PRT;     180 AA.
AC   P57771;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Regulator of G-protein signaling 8 (RGS8).
GN   RGS8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21218927; PubMed=11318611;
```

```
RA   Sood R., Bonner T.I., Malakowska I., Stephan D.A., Robbins C.M.,
RA   Connors T.D., Morgenbesser S.D., Su K., Faruque M.U., Pinkett H.,
RA   Graham C., Baxevanis A.D., Klinger K.W., Landes G.M., Trent J.M.,
RA   Carpten J.D.;
RT   "Cloning and characterization of 13 novel transcripts and the human
RT   RGS8 gene from the 1q25 region encompassing the hereditary prostate
RT   cancer (HPC1) locus.";
RL   Genomics 73:211-222(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
RT   "cDNA clones of human proteins involved in signal transduction
RT   sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. PREFERENTIALLY BINDS TO G(O)-ALPHA
CC       AND G(I)-ALPHA-3 (BY SIMILARITY).
-!- SIMILARITY: Contains 1 RGS domain.

CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).

DR   EMBL; AF297015; AAG45337.1; -.
DR   EMBL; AF300649; AAG18443.1; -.
DR   HSSP; P49799; 1AGR.
DR   Genew; HGNC:16810; RGS8.
DR   MIM; 607189; -.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor.
FT   DOMAIN       56    171       RGS.
FT   CONFLICT      1      9       MAALLMPRR -> MWNTLTRSLSDHPVGKDPQAMRTGQRQ
FT                               (IN REF. 1).
SQ   SEQUENCE   180 AA;  20917 MW;  198965B4C27F64C9 CRC64;

Query Match             33.6%;  Score 407.5;  DB 1;  Length 180;
Best Local Similarity   47.7%;  Pred. No. 5.2e-24;
Matches   83;  Conservative   32;  Mismatches   50;  Indels    9;  Gaps    2;

Qy    40  RAKEKRNRLSLLLQRPDFHGETQASRSALLAKE------TRVSPEBAVKWAESFDKLLSHR  94
          |||  |||||   ||        |||     ||||       ||   |||| |||||||
Db     9  RNKGMRTRLGCLSHKSD----SCSDFTAILPDKPNRALKRLSTEEATRWADSFDVLLSHK  64

Qy    95  DGVDAFTRFLKTEFSENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNID  154
          ||||| | ||||||||| ||||| ||||      ||| |||| |  |    |||||||
Db    65  YGVAAFRAFLKTEFSEENLEFWLACEEFKKTRSTAKIVSKAHRIFEEFVDVQAPREVNID  124

Qy    155 FHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEGRPQR  208
          | || |    |  || | |||||| | |||| |||||||||  ||| |  | |
Db    125 FQTREATRKNLQEPSLTCFDQAQGKVHSLMEKDSYPRFLRSKMYLDLLSQSQRR  178


RESULT 9
ID   RGS8_RAT                STANDARD;      PRT;     180 AA.
AC   P49804;
DT   01-OCT-1996 (Rel. 34, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Regulator of G-protein signaling 8 (RGS8).
GN   RGS8.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND FUNCTION.
RC   TISSUE=Hippocampus.
RX   MEDLINE=98054153; PubMed=9394004;
RA   Saitoh O., Kubo Y., Miyatani Y., Asano T., Nakata H.;
RT   "RGS8 accelerates G-protein-mediated modulation of K+ currents.";
RL   Nature 390:525-528(1997).
RN   [2]
RP   SEQUENCE OF 87-153 FROM N.A.
RC   TISSUE=Brain.
RX   MEDLINE=96140645; PubMed=8548815;
RA   Koelle M.R., Horvitz H.R.;
RT   "EGL-10 regulates G protein signaling in the C. elegans nervous
RT   system and shares a conserved domain with many mammalian proteins.";
RL   Cell 84:115-125(1996).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC   ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC   THEIR INACTIVE GDP-BOUND FORM. PREFERENTIALLY BINDS TO G(O)-ALPHA
CC   AND G(I)-ALPHA-3.
CC   -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN BRAIN. VERY
CC   LITTLE EXPRESSION DETECTED IN OTHER TISSUES.
CC   -!- DEVELOPMENTAL STAGE: DETECTED IN 13-DAY OLD EMBYOS. EXPRESSION
CC   INCREASES GRADUALLY IN LATER EMBRYOS AND MARKEDLY IN NEONATES TO
CC   ADULTS.
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AB006013; BAA23680.1; -.
DR   EMBL; U32432; AAC52369.1; -.
DR   HSSP; P49799; 1AGR.
DR   InterPro; IPR000342; Reg1_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Reg1_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor.
FT   DOMAIN   56   171   RGS.
SQ   SEQUENCE   180 AA; 20949 MW; 00FC35B4C278E857 CRC64;

Query Match   33.6%; Score 407.5; DB 1; Length 180;
Best Local Similarity   47.7%; Pred. No. 5.2e-24;
Matches   83; Conservative   32; Mismatches   50; Indels   9; Gaps   2;

Qy   40 RAKEKRNRLSLLQRPDFHGETQASRSAILAKE----TRVSPEEAVKWAESFDKLLSHR   94
                || ||||| : | | : : : :|| | ||| | | | |||:| |||
Db   9 RNKGMRTRLGCLSHKSD---SCSDFTAILPDKPNRALKRLSTEEATRWADSFDVLLSHK   64

Qy   95 DGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNID 154
             | |||| ||||||| ||||||| :: :| :   ::: | :|:| |:|:|||:|||
Db   65 YGVAAFRAFLKTEFSEENLEFWLACEEFKKTRSTAKLVTKAHRIFEEFVDVQAPREVNID 124

Qy   155 FHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEGRPQR 208
             |: :| :||| | | | | : :: | : : ||| || : | |||||
Db   125 FQTRRATRKNMQEPSLTCFDQAQGKVHSLMEKDSYPRFLRSKMYLDLLSQSQRR 178

RESULT 10
RGS5_MOUSE
ID   RGS5_MOUSE   STANDARD;   PRT;   181 AA.
AC   O08850; Q9DO22;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)

DE   Regulator of G-protein signaling 5 (RGS5).
GN   RGS5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97236828; PubMed=9079700;
RA   Chen C., Zheng B., Han J., Lin S.C.;
RT   "Characterization of a novel mammalian RGS protein that binds to
RT   Galpha proteins and inhibits pheromone signaling in yeast.";
RL   J. Biol. Chem. 272:8679-8685(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC   ACTIVITY OF G PROTEIN ALPHA SUBUNITS. THEREBY DRIVING THEM INTO
CC   THEIR INACTIVE GDP-BOUND FORM. BINDS TO G(I)-ALPHA AND G(O)-ALPHA
CC   BUT NOT TO G(S)-ALPHA.
CC   -!- TISSUE SPECIFICITY: Expressed in heart and muscle.
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; U67188; AAB50618.1; -.
DR   EMBL; AK004165; BAB23201.1; -.
DR   EMBL; BC037683; AAH37683.1; -.
DR   HSSP; P49799; 1AGR.
DR   MGD; MGI:1098434; Rgs5.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
KW   Signal transduction inhibitor.
FT   DOMAIN        64    180       RGS.
FT   CONFLICT      49     50       KA -> NG (IN REF. 1).
FT   CONFLICT      77     77       S -> T (IN REF. 1).
SQ   SEQUENCE   181 AA;  21085 MW;  B4B561CFE3DA9630 CRC64;

Query Match             33.5%;  Score 406.5;  DB 1;  Length 181;
Best Local Similarity   48.3%;  Pred. No. 6.3e-24;
Matches    85;  Conservative    29;  Mismatches    41;  Indels    21;  Gaps    2;

Qy     40  RAKEKRNRLSLLQRPDFHGETQASRSAL----------LAKETRVSPEEAVKNAESF  87
               ||||||:|||||||:| |||:|
Db     15  RAKEIKIKLGILLQKPD---------SAVDLVIPYNEKPEKAPKAKAHKPSLEEVLQWRQSL  65

Qy     88  DKLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQILLAKAIYEKFIQNDA  147
               |||:|::|:|||||::|||||||::||||||||||:||:||||||||||
Db     66  DKLLQNSYGFASFKSFLKSEFSEENLEFWVACENYKKIKSPIKVAEKAKQIYEEFIQTEA  125

Qy    148  PKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVQLMEHDSYKRFLKSETYLHLIE  203
               |||||||:||:|| |||:|:|:|||||:||:||:||||:|||||:||:|:
Db    126  PKEVNIDHFTKDITMKNLVEFSPRSFDLAQKRIYALMEKDSLPRFVRSEFYKELIK  181


RESULT 11
RGS4_HUMAN
ID   RGS4_HUMAN     STANDARD;      PRT;   205 AA.
AC   P49798;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Regulator of G-protein signaling 4 (RGS4) (RGP4).
GN   RGS4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96178495; PubMed=8602223;
RA   Druey K.M., Blumer K.J., Kang V.H., Kehrl J.H.;
RT   "Inhibition of G-protein-mediated MAP kinase activation by a new
RT   mammalian gene family.";
RL   Nature 379:742-746(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain, and Uterus;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Matsina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   PALMITOYLATION.
RX   MEDLINE=20076507; PubMed=10608901;
RA   Tu Y., Popov S., Slaughter C., Ross E.M.;
RT   "Palmitoylation of a conserved cysteine in the regulator of G protein
RT   signaling (RGS) domain modulates the GTPase-activating activity of
RT   RGS4 and RGS10.";
RL   J. Biol. Chem. 274:38260-38267(1999).
RN   [5]
RP   INHIBITION.
RX   MEDLINE=98421527; PubMed=9748280;
RA   Wang J., Ducret A., Tu Y., Kozasa T., Aebersold R., Ross E.M.;
RT   "RGSZ1, a Gz-selective RGS protein in brain. Structure, membrane
RT   association, regulation by Galphaz phosphorylation, and relationship
RT   to a Gz GTPase-activating protein subfamily.";
RL   J. Biol. Chem. 273:26014-26025(1998).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. ACTIVITY ON G(Z)-ALPHA IS INHIBITED
CC       BY PHOSPHORYLATION OF THE G-PROTEIN. ACTIVITY ON G(Z)-ALPHA AND
CC       G(I)-ALPHA-1 IS INHIBITED BY PALMITOYLATION OF THE G-PROTEIN.
CC   -!- PTM: Either Cys-2 or Cys-12 or both are palmitoylated.
CC   -!- PTM: Phosphorylated by cyclic GMP-dependent protein kinase (By
CC       similarity)
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; U27768; AAC50395.1; -.
DR   EMBL; AF493928; AAM12642.1; -.
DR   EMBL; BC000737; AAH00737.1; -.
DR   EMBL; BC051869; AAH51869.1; -.
DR   PIR; S78221; S78221.
DR   HSSP; P49799; 1AGR.
DR   Genew; HGNC:10000; RGS4.
DR   MIM; 602516; -.
DR   GO; GO:0005516; F:calmodulin binding; TAS.
DR   GO; GO:0005096; F:GTPase activator activity; TAS.
DR   GO; GO:0000188; P:inactivation of MAPK; TAS.
DR   GO; GO:0008277; P:regulation of G-protein coupled receptor pr...; TAS.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Lipoprotein; Palmitate;
KW   Phosphorylation.
FT   DOMAIN        62    178       RGS.
FT   LIPID          2      2       S-palmitoyl cysteine (Probable).
FT   LIPID         12     12       S-palmitoyl cysteine (Probable).
FT   LIPID         95     95       S-palmitoyl cysteine (Probable).
```

SQ   SEQUENCE   205 AA;   23255 MW;   7713F1F7496A698B CRC64;

Query Match   32.7%;   Score 397.5; DB 1; Length 205;
Best Local Similarity   47.0%;   Pred. No. 3.5e-23;
Matches   77; Conservative   35; Mismatches   49; Indels   3; Gaps   2;

Qy   41 AKEKRNRLSLLLQRPDF--HGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVD 98
               |:..::||:.|||:| ||||:|||:|:::||:|:|| |:||::|||||||
Db   16 AKDMKHRLGFLLQKSDSCEHSSSHSKKDKVVTCQ-RVSQEVKKWAESLENLISHECGLA 74

Qy   99 AFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTK 158
               ||:.|||:||||||:|:|:|||| ::|||::||||:||:|:|||||:|||
Db   75 AFKAFLKSKYSEENIDFWISCEEYKKIKSPSKLSPKAKKIYNEFISVQATKEVNLDSCTR 134

Qy   159 EVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLI 202
               ::|:||| .:: ||.| :: ||:| ||||||||:||||:|:
Db   135 EETSRNMLEPTITCFDEAQKKIFNLMEKDSYRRFLKSRFYLDLV 178

RESULT 12
RGS4_MOUSE
ID   RGS4_MOUSE   STANDARD;   PRT;   205 AA.
AC   O08899; Q99L30;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Regulator of G-protein signaling 4 (RGS4).
GN   RGS4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Substantia nigra;
RX   MEDLINE=98086343; PubMed=9425263;
RA   Nomoto S., Adachi K., Yang L.X., Hirata Y., Muraguchi S., Kiuchi K.;
RT   "Distribution of RGS4 mRNA in mouse brain shown by in situ
RT   hybridization.";
RL   Biochem. Biophys. Res. Commun. 241:281-287(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast tumor;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC   ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC   THEIR INACTIVE GDP-BOUND FORM. ACTIVITY ON G(Z)-ALPHA IS INHIBITED
CC   BY PHOSPHORYLATION OF THE G-PROTEIN. ACTIVITY ON G(Z)-ALPHA AND
CC   G(I)-ALPHA-1 IS INHIBITED BY PALMITOYLATION OF THE G-PROTEIN (BY
CC   SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN BRAIN, MODERATELY
CC   LOW LEVELS IN HEART, AND VERY LOW LEVELS IN LUNG, LIVER, AND
CC   SKELETAL MUSCLE.
CC   -!- PTM: Either Cys-2 or Cys-12 or both are palmitoylated (By
CC   similarity).
CC   -!- PTM: Phosphorylated by cyclic GMP-dependent protein kinase (By
CC   similarity).
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AB004315; BAA20400.1; -.
DR   EMBL; BC003882; AAH03882.1; -.
DR   HSSP; P49799; 1AGR.
DR   MGD; MGI:108409; Rgs4.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Lipoprotein; Palmitate;
KW   Phosphorylation.
FT   DOMAIN   62   178   RGS.
FT   LIPID   2   2   S-palmitoyl cysteine (By similarity).
FT   LIPID   12   12   S-palmitoyl cysteine (By similarity).
FT   LIPID   95   95   S-palmitoyl cysteine (By similarity).
FT   CONFLICT   162   162   R -> K (IN REF. 2).
SQ   SEQUENCE   205 AA;   23288 MW;   5D795817I1A1F67C CRC64;

Query Match   32.7%;   Score 397.5; DB 1; Length 205;
Best Local Similarity   47.2%;   Pred. No. 3.5e-23;
Matches   77; Conservative   34; Mismatches   49; Indels   3; Gaps   2;

Qy   41 AKEKRNRLSLLLQRPDF--HGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVD 98
               |:..::||:.|||:| ||||:|||:|:::||:|:|| |:||::|||||||
Db   16 AKDMKHRLGFLLQKSDSCEHSSSHSKKDKVVTCQ-RVSQEVKKWAESLENLISHECGLA 74

Qy   99 AFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTK 158
               ||:.|||:||||||:|:|:|||| ::|||::||||:||:|:|||||:|||
Db   75 AFKAFLKSKYSEENIDFWISCEEYKKIKSPSKLSPKAKKIYNEFISVQATKEVNLDSCTR 134

Qy   159 EVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHL 201
               ::|:||| .:: ||.| :: ||:| ||||||||:||||:|
Db   135 EETSRNMLQPTITCFDEAQKKIFNLMERDSYRRFLKSRFYLDL 177

RESULT 13
RGS3_MOUSE
ID   RGS3_MOUSE   STANDARD;   PRT;   568 AA.
AC   Q9DC04; Q9JL22; Q9JL23;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Regulator of G-protein signaling 3 (RGS3).
GN   RGS3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RC   STRAIN=C57BL/6J; TISSUE=Lung;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477993;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   PARTIAL SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=20243574; PubMed=10779778;
RA   Reif K., Cyster J.G.;
RT   "RGS molecule expression in murine B lymphocytes and ability to
RT   down-regulate chemotaxis to lymphoid chemokines.";
RL   J. Immunol. 164:4720-4729(2000).
CC   -!- FUNCTION: Inhibits signal transduction by increasing the GTPase
CC   activity of G protein alpha subunits thereby driving them into
CC   their inactive GDP-bound form.
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=2;
CC   Name=Long;
CC   IsoId=Q9DC04-1; Sequence=Displayed;
CC   Name=Short;
CC   IsoId=Q9DC04-2; Sequence=VSP_005663, VSP_005664;
CC   -!- PTM: Phosphorylated by cyclic GMP-dependent protein kinase (By
CC   similarity).
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AK004648; BAB23439.1; -.
DR   EMBL; BC033449; AAH33449.1; -.
DR   EMBL; AF156670; AAF34627.1; -.
DR   EMBL; AF215669; AAF34626.1; -.
DR   HSSP; P49799; 1AGR.
DR   MGD; MGI:1354734; Rgs3.

DR   GO; GO:0005634; C:nucleus; IDA.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50112; RGS; 1.
KW   Signal transduction inhibitor; Alternative splicing; Phosphorylation.
FT   DOMAIN   443   559   RGS.
FT   VARSPLIC   1   375   Missing (in isoform Short).
FT   /FTId=VSP_005663.
FT   VARSPLIC   376   397   KMSGTDITECDEASRKRKSKNI -> MLRGMYLTRNGNLQR
FT   RHTMKE (in isoform Short).
FT   /FTId=VSP_005664.
SQ   SEQUENCE   568 AA;   61558 MW; 91939E22F3DA9F5C CRC64;

Query Match   32.6%; Score 396; DB 1; Length 568;
Best Local Similarity   44.2%; Pred. No. 1.5e-22;
  Matches   80; Conservative   38; Mismatches   61; Indels   2; Gaps   1;

Qy   31 EETSIEAKIR--AKEKRNRLSLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFD 88
          ||:||.||:| ::|:|:::|:.::::|| :|:||:||:||||||||||||::||:|
Db   386 DEASRRKSKNIAKDMQNKLAIFRRNESPGAQPASKTDKTTKSFKPTSEEALKWSESLE 445

Qy   89 KLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAP 148
          ||:::::||.||||:||||||||||||||||:|||||||||||||.|||||||||||||
Db   446 KLLLHKYGLEVFQAFLRTEFSEENLEFWLACEDFKKVKSQSKMAAKAKKIFAEFIAIQAC 505

Qy   149 KEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEGRPQR 208
          ||||||:|||:| ::|.:|::||:::|||||||||||||||||||||:||||||.|| :
Db   506 KEVNLDSYTREHTKENLQSITRGCFDLAQKRIFGLMEKDSYPRFLRSDLIYLDLINQKKMS 565

Qy   209 P 209
          |
Db   566 P 566

RESULT 14
RGS4_RAT
ID   RGS4_RAT   STANDARD;   PRT;   205 AA.
AC   P49799;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Regulator of G-protein signaling 4 (RGS4) (RGP4).
GN   RGS4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96178495; PubMed=8602223;
RA   Druey K.M., Blumer K.J., Kang V.H., Kehrl J.H.;
RT   "Inhibition of G-protein-mediated MAP kinase activation by a new
RT   mammalian gene family.";
RL   Nature 379:742-746(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RA   Zhou M.-Y., Gomez-Sanchez C.E., Gomez-Sanchez E.P.;
RT   "The complete cDNA sequence analysis of the rat RGS4.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 93-159 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=96140645; PubMed=8548815;
RA   Koelle M.R., Horvitz H.R.;
RT   "EGL-10 regulates G protein signaling in the C. elegans nervous
RT   system and shares a conserved domain with many mammalian proteins.";
RL   Cell 84:115-125(1996).
RN   [4]
RP   PHOSPHORYLATION.

```
RX   MEDLINE=20167219; PubMed=10702309;
RA   Pedram A., Razandi M., Kehrl J., Levin E.R.;
RT   "Natriuretic peptides inhibit G protein activation. Mediation through
RT   cross-talk between cyclic GMP-dependent protein kinase and regulators
RT   of G protein-signaling proteins.";
RL   J. Biol. Chem. 275:7365-7372(2000).
RN   [5]
RP   INHIBITION.
RX   MEDLINE=98016286; PubMed=9351196;
RA   Tu Y., Wang J., Ross E.M.;
RT   "Inhibition of brain Gz GAP and other RGS proteins by palmitoylation
RT   of G protein alpha subunits.";
RL   Science 278:1132-1135(1997).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF COMPLEX WITH G(I)-ALPHA.
RX   MEDLINE=97726066; PubMed=9108480;
RA   Tesmer J.J.G., Berman D.M., Gilman A.G., Sprang S.R.;
RT   "Structure of RGS4 bound to AlF4-activated G(i alpha):
RT   stabilization of the transition state for GTP hydrolysis.";
RL   Cell 89:251-261(1997).
CC   -!- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC       ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC       THEIR INACTIVE GDP-BOUND FORM. ACTIVITY ON G(Z)-ALPHA IS INHIBITED
CC       BY PHOSPHORYLATION OF THE G-PROTEIN. ACTIVITY ON G(Z)-ALPHA AND
CC       G(I)-ALPHA-1 IS INHIBITED BY PALMITOYLATION OF THE G-PROTEIN.
CC   -!- PTM: Either Cys-2 or Cys-12 or both are palmitoylated (By
CC       similarity).
CC   -!- PTM: Phosphorylated by cyclic GMP-dependent protein kinase.
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; U27767; AAC52440.1; -.
DR   EMBL; AF17211; AAD12065.1; -.
DR   EMBL; U32327; AAC52367.1; -.
DR   PDB; 1AGR; 16-JUN-97.
DR   PDB; 1EZT; 15-JAN-01.
DR   PDB; 1EZY; 15-JAN-01.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Lipoprotein; Palmitate;
KW   Phosphorylation; 3D-structure.
FT   DOMAIN       62    178       RGS.
FT   LIPID         2      2       S-palmitoyl cysteine (By similarity).
FT   LIPID        12     12       S-palmitoyl cysteine (By similarity).
FT   LIPID        95     95       S-palmitoyl cysteine (By similarity).
FT   HELIX        53     59
FT   TURN         60     61
FT   HELIX        63     68
FT   HELIX        70     82
FT   TURN         83     84
FT   TURN         86     86
FT   HELIX        87    100
FT   TURN        104    106
FT   HELIX       107    118
FT   HELIX       119    119
FT   TURN        121    122
FT   TURN        124    125
FT   HELIX       131    140
FT   TURN        141    142
FT   TURN        146    149
FT   HELIX       150    162
FT   TURN        163    163
```

```
FT   HELIX       164    168
FT   TURN        169    170
FT   HELIX       172    175
FT   TURN        176    176
SQ   SEQUENCE    205 AA;  23248 MW;  9647C0EC909D0F6F CRC64;

Query Match             32.6%;  Score 395.5;  DB 1;  Length 205;
Best Local Similarity   46.6%;  Pred. No. 4.9e-23;
Matches   76;  Conservative  36;  Mismatches  48;  Indels   3;  Gaps   2;

Qy     41 AKEKRNRLSLLQRPDF--HGETQASRGALLAKETRVSPEEAVKWAESFDKLLSHRDGVD  98
          |||:::||::|||:||::||:|:|||::||:|::|||||:|  |||:|:|  ::||||
Db     16 AKDMKHRLGFLLQKSDSCEHSSSHSKKDKVVTCQ-RVSQEEVKKWAESLENLINHECGLA  74

Qy     99 AFTRFLKTEFSEENIEFWVACEDFKKCKEPQQILKAKAIYEKFIQNDAPKEVNIDFHTK 158
          ||||::||::||::||:|::|:::|||::|||:||:|:|:||  ::||:   |:|
Db     75 AFKAFLKSEYSEENIDFWISCEEYKKIKSPSKLSPKAKKIYNEFISVQATKEVNLDSCTR 134

Qy    159 EVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHL 201
          ||:||:||::|||:|:|::|:|:::|:||:||:||||:||:||
Db    135 EETSRNMLEPTITCFDEAQKKIFNLMEKDSYRRFLKSRFYLDL 177


RESULT 15
ID   RGS3_HUMAN     STANDARD;      PRT;    519 AA.
AC   P49796; Q8TD59; Q8TD68;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Regulator of G-protein signaling 3 (RGS3) (RGP3).
GN   RGS3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=96178495; PubMed=8602223;
RA   Druey K.M., Blumer K.J., Kang V.H., Kehrl J.H.;
RT   "Inhibition of G-protein-mediated MAP kinase activation by a new
RT   mammalian gene family.";
RL   Nature 379:742-746(1996).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Brain;
RA   Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
RT   "cDNA clones of human proteins involved in signal transduction
RT   sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Skuce C.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Inhibits signal transduction by increasing the GTPase
CC       activity of G protein alpha subunits thereby driving them into
CC       their inactive GDP-bound form.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P49796-1; Sequence=Displayed;
CC       Name=2; Synonyms=RGS3T;
CC         IsoId=P49796-2; Sequence=VSP_005662;
CC   -!- PTM: Phosphorylated by cyclic GMP-dependent protein kinase (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 RGS domain.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------
DR   EMBL; U27655; AAC50394.1; -.
DR   EMBL; AF493927; AAM12641.1; -.
DR   EMBL; AF493941; AAM12655.1; -.
DR   EMBL; AL162727; CAC78977.1; -.
DR   PIR; S78089; S78089.
DR   HSSP; P49799; 1AGR.
DR   Genew; HGNC:9999; RGS3.
DR   MIM; 602189; -.
DR   GO; GO:0005829; C:cytosol; TAS.
DR   GO; GO:0005096; F:GTPase activator activity; TAS.
DR   GO; GO:0000188; P:inactivation of MAPK; TAS.
DR   GO; GO:0008277; P:regulation of G-protein coupled receptor pr. . .; TAS.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Signal transduction inhibitor; Alternative splicing; Phosphorylation.
FT   DOMAIN      394    510       RGS.
FT   VARSPLIC      1    313       Missing (in isoform 2).
FT                                /FTId=VSP_005662.
FT   CONFLICT    305    305       K -> R (IN REF. 2).
SQ   SEQUENCE   519 AA;  56601 MW;  F1CFE3F27D4673A0 CRC64;

Query Match             31.5%;  Score 383;  DB 1;  Length 519;
Best Local Similarity   43.7%;  Pred. No. 1.3e-21;
Matches    76;  Conservative    35;  Mismatches    63;  Indels     0;  Gaps     0;

Qy       36 EAKIRAKEKRNRLSLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRD 95
               ::| |||:|:|:||   |             :|: :| ||||| ||| :|||||
Db      344 KSKNLAKDMKNKLGGIFRRRNESPGAPPAGKADKMMKSFKPTSEEALKWGESLEKLLVHKY 403

Qy       96 GVDAFTRFLKTEFSEENIBFWVACEDFKKCKBPQQIILKAKAIYEKFIQNDAPKEVNIDF 155
             |:  | ||||||:||:|||:|  :: |||||   | | ||||| | |||||:
Db      404 GLAVFQAFLRTEFSEENLEFWLACEDFKKVKSQSKVASKAKKIFAEYIAIQACKEVNLDS 463

Qy      156 HTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEGRPQRP 209
             |:|  :: :: .: ||||| ||:|||||:||:||||||:|:||:||| :: :|
Db      464 YTREHTKDNLQSVTRGCFDLAQKRIFGLMEKDSYPRFLRSDLYLDLINQKKMSP 517
```

```
Search completed: August 20, 2004, 16:41:30
Job time : 15.5 secs
```

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:        August 20, 2004, 15:52:01 ; Search time 74.5 Seconds
                                          (without alignments)
                                          995.259 Million cell updates/sec

Title:         US-09-894-749-4
Perfect score: 1214
Sequence:      1 MDMSLVFFSQLNMCESKEKT.........SRSFTYNDFQDVKSDVAIWL 235

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:       1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    SPTREMBL_25:*
              1: sp_archea:*
              2: sp_bacteria:*
              3: sp_fungi:*
              4: sp_human:*
              5: sp_invertebrate:*
              6: sp_mammal:*
              7: sp_mhc:*
              8: sp_organelle:*
              9: sp_phage:*
             10: sp_plant:*
             11: sp_rodent:*
             12: sp_virus:*
             13: sp_vertebrate:*
             14: sp_unclassified:*
             15: sp_rvirus:*
             16: sp_bacteriap:*
             17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                          SUMMARIES
```

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	421	34.7	213	13	Q7ZZS4	Q7zzs4 gallus gall
2	421	34.7	225	13	Q7ZZS5	Q7zzs5 gallus gall
3	411.5	33.9	164	6	Q8HXV9	Q8hxv9 ovis aries
4	410.5	33.8	180	11	Q8BXT1	Q8bxt1 mus musculu
5	406.5	33.5	181	6	Q86422	Q86422 sus scrofa
6	404.5	33.4	196	6	Q95K68	Q95k68 macaca fasc
7	404	33.3	208	13	Q7SZC6	Q7szc6 gallus gall
8	400.5	33.0	182	13	Q7T2D3	Q7t2d3 brachydanio
9	397	32.7	283	13	Q7SYI2	Q7syi2 gallus gall
10	397	32.7	408	13	Q7SYI1	Q7syi1 gallus gall
11	397	32.7	441	13	Q7SYI0	Q7syi0 gallus gall
12	397	32.7	799	13	Q7SYH9	Q7syh9 gallus gall
13	396	32.6	930	11	Q92SG9	Q92sg9 mus musculu
14	391	32.2	967	11	Q92QQ9	Q92qq9 rattus norv
15	384	31.6	319	4	Q8NFN6	Q8nfn6 homo sapien
16	383	31.5	917	4	Q8NFN4	Q8nfn4 homo sapien
17	383	31.5	917	4	Q8WXA0	Q8wxa0 homo sapien
18	383	31.5	917	4	Q8IUQ1	Q8iuq1 homo sapien
19	383	31.5	1093	4	Q8NFN5	Q8nfn5 homo sapien
20	382	31.5	192	4	Q8WVE9	Q8wve9 homo sapien
21	381	31.4	192	4	Q8WV02	Q8wv02 homo sapien
22	380	31.3	284	4	Q96NV5	Q96nv5 homo sapien
23	377	31.1	168	13	Q7SYI3	Q7syi3 gallus gall
24	356	29.3	235	11	Q80V16	Q80v16 mus musculu
25	355.5	29.3	204	13	Q19D9	Q19d9 xenopus lae
26	352.5	29.0	201	11	Q7TNU9	Q7tnu9 mus musculu
27	309	25.5	271	5	Q9V888	Q9v888 drosophila
28	309	25.5	274	5	Q8T017	Q8t017 drosophila
29	306.5	25.2	194	11	Q8BFU4	Q8bfu4 mus musculu
30	304	25.0	158	11	Q8K443	Q8k443 mus musculu
31	303.5	25.0	194	11	Q8CGT5	Q8cgt5 mus musculu
32	301.5	24.8	220	13	Q7TOQ6	Q7t0q6 xenopus lae
33	300.5	24.8	220	13	Q8QHK0	Q8qhk0 xenopus lae
34	300	24.7	208	11	Q8C5J7	Q8c5j7 mus musculu
35	297.5	24.5	199	13	Q8AYF2	Q8ayf2 gallus gall
36	297	24.5	210	11	Q8C5F3	Q8c5f3 mus musculu
37	296	24.4	230	11	Q8BR34	Q8br34 mus musculu
38	294	24.2	150	4	Q8TD61	Q8td61 homo sapien
39	291.5	24.0	181	5	Q9TVK0	Q9tvk0 caenorhabdi
40	290	23.9	244	5	O96842	O96842 drosophila
41	280.5	23.1	181	4	Q9GGN0	Q9ggn0 homo sapien
42	273.5	22.5	424	5	Q95Q48	Q95q48 caenorhabdi
43	273.5	22.5	475	4	Q724K3	Q724k3 homo sapien
44	273.5	22.5	477	4	Q724K4	Q724k4 homo sapien
45	273.5	22.5	480	4	Q724K5	Q724k5 homo sapien

```
                          ALIGNMENTS

RESULT 1
Q7ZZS4
ID   Q7ZZS4      PRELIMINARY;      PRT;     213 AA.
AC   Q7ZZS4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G-protein signaling protein 2.
GN   RGS2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tosetti P., Parente V., Taglietti V., Dunlap K., Toselli M.;
RT   "Concentration-dependent effects of chick RGS2L on neuronal L-type Ca
RT   channel modulation by bradykinin.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF502149; AAP30802.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE  213 AA;  24786 MW;  B0CAD39ACC496D1E CRC64;

Query Match             34.7%;  Score 421; DB 13;  Length 213;
Best Local Similarity   50.3%;  Pred. No. 1.4e-25;
Matches   83; Conservative   31; Mismatches   51; Indels   0; Gaps   0;

Qy     42 KEKRNRLSLLQRPDFHGETQASRSALLAKETRVSPEAVKWAESFDKLLSHRDGVDAFT 101
          ||||  |||:  | ||  | |:::::    || | ||:  || :|:| ||| | ||
Db     41 KDWKKKLSYFLQNSSRSSKVKSTKAGKHTYFRPSPEZARLWSEAFDELLANKYGVAAFR 100

Qy    102 RFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVI 161
```

Db   101 AFLKSEFCEENIEFWLACEDFKKTKSPQKLTLKAKKIYNDFIEKEAPKEINIDFQTKNMI 160
            |||:||| |||||||||:||||||| ||:||| |||||||||||||||||||||:|||||:|

Qy   162 AKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEGRP 206
            |:::.|| | |::|:|||| |||||::|

Db   161 AQNLQEATHTCFSAAQKRVYSLMENNSYPRFLESEFYQELCKKPP 205

RESULT 2
Q7ZZS5
ID   Q7ZZS5;
AC   Q7ZZS5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G-protein signaling protein 2 long isoform.
GN   RGS2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tosetti P., Parente V., Taglietti V., Dunlap K., Toselli M.;
RT   "Concentration-dependent effects of chick RGS2L on neuronal L-type Ca
RT   channel modulation by bradykinin.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF502148; AAP30801.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE   225 AA; 25956 MW; 5F8C5FAE2C665EE1 CRC64;

Query Match   34.7%; Score 421; DB 13; Length 225;
Best Local Similarity   50.3%; Pred. No. 1.5e-25;
Matches   83; Conservative   31; Mismatches   51; Indels   0; Gaps   0;

Qy   42 KEKRNRLSLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFT 101
            ||::||| | |::|:|||| |||||::|||||||||||||||||||||||||:|

Db   53 KDWKKKLSYFLQNSSRSSKVKSTKAGKHHTYFRPSPEEARLWSEAFDELLANKYGVAAFR 112

Qy   102 RFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVI 161
            |||||:|||||||||:|||||||:|||||||||||:|||||:||:||||||||||:||:|

Db   113 AFLKSEFCEENIEFWLACEDFKKTKSPQKLTLKAKKIYNDFIEKEAPKEINIDFQTKNMI 172

Qy   162 AKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEGRP 206
            |:::.|| | |::|:|||| |||||::|

Db   173 AQNLQEATHTCFSAAQKRVYSLMENNSYPRFLESEFYQELCKKPP 217

RESULT 3
Q8HXV9
ID   Q8HXV9;   PRELIMINARY;   PRT;   164 AA.
AC   Q8HXV9;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   RGS2 (Fragment).
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Airway;
RA   Nlend M.C., Bookman R.J., Conner G.E., Salathe M.;
RT   "RGS2 Modulates Purinergic Calcium and Ciliary Beat Frequency

---

RT   Responses in Airway Epithelia.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF472577; AAN32893.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
FT   NON_TER   1   1
FT   NON_TER   164   164
SQ   SEQUENCE   164 AA; 18947 MW; EB5B828CB05E43D5 CRC64;

Query Match   33.9%; Score 411.5; DB 6; Length 164;
Best Local Similarity   51.2%; Pred. No. 5.8e-25;
Matches   84; Conservative   28; Mismatches   51; Indels   1; Gaps   1;

Qy   47 RLSLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAFTRFLKT 106
            ||||||:|:|::|:|||| |||||::|:||:|::::||:||:|||::|||::|::|

Db   2 RLSYFLQNSSSPGKPTGKKSKQQTFKPSPEEAQLWSEAFDELLASKYGLAAFRAFLKS 61

Qy   107 EFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEVIAKSIA 166
            |:||||||||:|||||||:|||||||||||:|||||:||:||||||||||:||:|:::.

Db   62 EYCEENIEFWLACEDFKKTKSPQKLSSKAKKIYTDFIEKEAPKEINIDFQTKTLIAQNIQ 121

Qy   167 QPPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEGRPQPT 210
            ::| |::|:|||| |||||::|||||||:||:|||||||

Db   122 EATSGCFTTAQKRVYSLMENNSYPRFLESEFYQDLCK-KPQITT 164

RESULT 4
Q8BXT1
ID   Q8BXT1;   PRELIMINARY;   PRT;   180 AA.
AC   Q8BXT1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Regulator of G-protein signaling 8.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Retina;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK044337; BAC31874.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE   180 AA; 20963 MW; 00FC35E572785856 CRC64;

Query Match   33.8%; Score 410.5; DB 11; Length 180;
Best Local Similarity   48.3%; Pred. No. 7.8e-25;
Matches   84; Conservative   31; Mismatches   50; Indels   9; Gaps   2;

Qy   40 RAKEKRNLSLLQRPDFHGETQASRSALLAKE-----TRVSPEEAVKWAESFDKLLSHR 94
            |:| |||:||||||||||:|||| |||:| ||||||||||:||||||

Db   9 RNKGMRTRLGCLSHKSD----SCSDFTAILPDKPNRALKRLSTEEATRWAESFDVLLSHK 64

Qy   95 DGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNID 154
            |||||:|::|::|:||||||||:|||||||:|||||||||||:|||||:

Db   65 YGVAAFRAFLKTEFSEENLEFWLACEEFKKTRSTAKLVTKAHRIFSEFVDVQAPREVNID 124

```
Qy  155 FHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEGRPQR 208
         |:::|| |:: :|| || ||||||||||||||||||:: |:: || |:::
Db  125 FQTREATRKNMQEPSLTCFDQAAQGKVHSLMEKDSYPRFLRSKMYLDLLSQSQRR 178


RESULT 5
Q84Z2
ID   Q84Z2       PRELIMINARY;       PRT;   181 AA.
AC   Q84Z2;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G-protein signalling 5.
GN   RGS5.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Aorta;
RA   Weiss B., Hehigans S.I., Wietelmann A.B., Zeyer A., Richter M.,
RA   Kloevekorn W.P., Zimmermann R., von der Ahe D.;
RT   "Differential gene expression in atherosclerotic coronary arteries.";
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ549925; CAD71257.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE  181 AA;  21054 MW;  BEB98C7A4A99D669 CRC64;

Query Match      33.5%;  Score 406.5;  DB 6;  Length 181;
Best Local Similarity   48.5%;  Pred. No. 1.6e-24;
Matches    81;  Conservative  34;  Mismatches  49;  Indels   3;  Gaps   1;

Qy   40 RAKEKRNRLSLLQRPDFHGETQASRSALLAKETRV---SPEEAVKWAESFDKLLSHRDG 96
         ||||:||||:||||:|:|:   |||::|||:|||||   :|||||||||:|||||||
Db   15 RAKEIKIKLGILLQKPESAVDLVIPYNEKPDKPVKIQKPSLDEALQWRDSLDKLLQNNYG 74

Qy   97 VDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQILKAKAIYEKFIQNDAPKEVNIDFH 156
         :|| ||||||||||||||||:| :: :||:||||||||||| ||:||
Db   75 LASFKSFLKSEFSEENLEFWMACSDYKKIKSPVKMAEKAKKIYEEFIQSEAPKEVNIDHF 134

Qy  157 TKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIE 203
         ||:||||||||||||||||||||||||||||||:|||||||| ||
Db  135 TKEITMKNLVEPSPSSFDVAQKRIYALMEKDSLPRFVRSEFYQEFIK 181


RESULT 6
Q95K68
ID   Q95K68      PRELIMINARY;       PRT;   196 AA.
AC   Q95K68;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Medulla oblongata;
RA   Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA   Suzuki Y., Sugano S., Hashimoto K.;
RT   "Isolation of full-length cDNA clones from macaque brain cDNA
RT   libraries.";
```

```
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 RGS DOMAIN.
DR   EMBL; AB066513; BAB62198.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
KW   Hypothetical protein.
SQ   SEQUENCE  196 AA;  22939 MW;  A42C6D41A244064 CRC64;

Query Match      33.3%;  Score 404.5;  DB 6;  Length 196;
Best Local Similarity   45.7%;  Pred. No. 2.6e-24;
Matches    85;  Conservative  29;  Mismatches  55;  Indels  17;  Gaps   2;

Qy   40 RAKEKRNRLSLLQRPDFHGETQA--------SRSAL-------------LAKETRVSPEEAVK 82
          |: |: : :|: |:   |||  :|:            :|||||| | :::|||:|:||:
Db    9 RNKGMRTLGCLSHKSDSCSDFTAILPDKPNRALNYLRMYKFTATELQESRRLSTERATR 68

Qy   83 WAESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQILKAKAIYEKF 142
         ||:||| |||: |  ||: ||||||||||:|||:|||:|||::|:| |:::||||||:|
Db   69 WADSFDVLLSHKYGVAAFRAFLKTEFSEENLEFWLACEEFKKTRSTAKLVSKAHRIFSEF 128

Qy  143 IQNDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHI 202
         : :|| :|||||| |:|  |: :|:||| || |: ||:|::|||:||:||| |:||| :
Db  129 VDVQAPREVNIDFQTREATRKNMQEPSLTCFDQAQGKVHSLMEKDSYPRFLRSKMYLDLL 188

Qy  203 EGRPQR 208
         : ||:
Db  189 SQSQRR 194


RESULT 7
Q7SZC6
ID   Q7SZC6      PRELIMINARY;       PRT;   208 AA.
AC   Q7SZC6;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G-protein signaling 4.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=P4 hindbrain;
RA   Grillet N., Dubreuil V., Dufour H., Goridis C., Brunet J.-F.;
RT   "Dynamic expression of RGS4 in the developing nervous system and
RT   control by the paired-like homeoprotein Phox2b.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY297457; AAP57222.1; -.
SQ   SEQUENCE  208 AA;  23523 MW;  838B31750F0EA68 CRC64;

Query Match      33.3%;  Score 404;  DB 13;  Length 208;
Best Local Similarity   47.2%;  Pred. No. 3e-24;
Matches    85;  Conservative  35;  Mismatches  50;  Indels  10;  Gaps   5;

Qy   41 AKEKRNRLSLLQRPDF--HGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVD 98
         ||: :||| :|||:|     |: ::|: | :|  ::||: :|| |||:: :|:|::|||
Db   16 AKDMKHRLGVLLQKSDSCDYGSSQGKKEXVSSSQ-RVSQEEVKKWAESLENLIHHDRGLA 74

Qy   99 AFTRFLKTEFSEENIEFWVACEDFKKCKEPQQILKAKAIYEKFIQNDAPKEVNIDFHTK 158
         ||: ||| | ||||||||| |||:||  | : :: || :||::|| | | ||||:| |:
Db   75 AFRAFLKSEYSEENIEFWVSCEDYKKTSPAKLSTKARKIYDEFISVQATKEVNLDSCTR 134

Qy  159 EVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEGRPQR----PTNLRR 214
         |  : :  ||| | || || ::: |: |||:||||| || ||:   |      ||: :|
Db  135 EKTSHNMLEPTLSCFDEAQRKIFTLMKDSYRRFLKS-PYLDLVS--PPRAGCGPENCKR 191
```

RESULT 8
```
ID   Q7T2D3          PRELIMINARY;      PRT;   182 AA.
AC   Q7T2D3;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC054594; AAH54594.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE  182 AA;  21053 MW;  4D502817AA8CE4FF CRC64;
```

```
Query Match             33.0%;  Score 400.5;  DB 13;  Length 182;
Best Local Similarity   47.5%;  Pred. No. 4.9e-24;
Matches   85;  Conservative  29;  Mismatches  52;  Indels  13;  Gaps   3;

Qy   32 ETSIEAKIRAKEKRNRLSLLQRPDF-------HGETQASRSALLAKETRVSPEEAVKWA 84
            ||||||  |||||| ::||  :| ||         | ||||||  |||||  ||||:
Db   10 QTCLE---RAKEIKTKLGTLLQKPENSIDLIIPYQEKPEKKPEKLQKAT--PEEAAQWR 63

Qy   85 ESFDKLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQ 144
           |||  |||||||||||||||||||||||||||||||| ||||||| ||||||||||||||
Db   64 ESLDKVLSNSYGLATFKSFLRSEFSEENIEFWEACEDFKKTKNPLKMATKAKKIYEDFIQ 123

Qy  145 NDAPKEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIE 203
            |||:||||||||| :: || ||||||:||||||||||||||| ||:|:  ||:||:|  |:
Db  124 TGGPKEVNIDHFTKDVTLRNLVDLSSSTFELAQSRIYTLMEKDSFGRFLRSDQYQELLK 182
```

RESULT 9
```
ID   Q7SYI2          PRELIMINARY;      PRT;   283 AA.
AC   Q7SYI2;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G protein signaling 3 RGS3ss isoform.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=DRG neuron;
RX   MEDLINE=22684492; PubMed=12771384;
RA   Tosetti P., Pathak N., Jacob M.H., Dunlap K.;
RT   "RGS3 mediates a calcium-dependent termination of G protein signaling
RT   in sensory neurons.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7337-7342(2003).
DR   EMBL; AY124774; AAM94020.1; -.
SQ   SEQUENCE  283 AA;  32070 MW;  5C9871357F00A195 CRC64;
```

```
Query Match             32.7%;  Score 397;  DB 13;  Length 283;
Best Local Similarity   45.4%;  Pred. No. 1.5e-23;
Matches   79;  Conservative  38;  Mismatches  55;  Indels  2;  Gaps   1;

Qy   31 EETSIEAKIR--AKEKRNRLSLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFD 88
           |||| :||  ||  |||||:||||||||| |||||||||||||||||||||  ||||||
Db  101 EETSRKKRSRSLAKDMNRLGIFRRRNESPGANPSKLDKVLKSLKPAPEEALKWGESLE 160

Qy   89 KLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAP 148
           |||||||||||||||||||||||||||||||||||  ::|:| |||||||||||||||  :
Db  161 KLLLHKYGLAAFRAFLRTEFSEENLEFWLACEEYKKIKSQSKMVSKAKKIFAEYIAIQSC 220

Qy  149 KEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLI 202
           ||||:|||||||| ::|  :|||||||||||||:|||  ||:|||:  || :||
Db  221 KEVNLDSYTREHTKENLQNITRGCFDLAQKRIYGLMEKDSYPRFLRSDLYLDII 274
```

RESULT 10
```
ID   Q7SYI1          PRELIMINARY;      PRT;   408 AA.
AC   Q7SYI1;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G protein signaling 3 RGS3s isoform.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=DRG neuron;
RX   MEDLINE=22684492; PubMed=12771384;
RA   Tosetti P., Pathak N., Jacob M.H., Dunlap K.;
RT   "RGS3 mediates a calcium-dependent termination of G protein signaling
RT   in sensory neurons.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7337-7342(2003).
DR   EMBL; AY124775; AAM94021.1; -.
SQ   SEQUENCE  408 AA;  45438 MW;  75B50794F22E74B5 CRC64;
```

```
Query Match             32.7%;  Score 397;  DB 13;  Length 408;
Best Local Similarity   45.4%;  Pred. No. 2.4e-23;
Matches   79;  Conservative  38;  Mismatches  55;  Indels  2;  Gaps   1;

Qy   31 EETSIEAKIR--AKEKRNRLSLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFD 88
           |||| :||  ||  |||||:||||||||| |||||||||||||||||||||  ||||||
Db  226 EETSRKKRSRSLAKDMNRLGIFRRRNESPGANPSKLDKVLKSLKPAPEEALKWGESLE 285

Qy   89 KLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAP 148
           |||||||||||||||||||||||||||||||||||  ::|:| |||||||||||||||  :
Db  286 KLLLHKYGLAAFRAFLRTEFSEENLEFWLACEEYKKIKSQSKMVSKAKKIFAEYIAIQSC 345

Qy  149 KEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLI 202
           ||||:|||||||| ::|  :|||||||||||||:|||  ||:|||:  || :||
Db  346 KEVNLDSYTREHTKENLQNITRGCFDLAQKRIYGLMEKDSYPRFLRSDLYLDII 399
```

RESULT 11
Q7SYI0
ID   Q7SYI0   PRELIMINARY;   PRT;   441 AA.
AC   Q7SYI0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G protein signaling 3 (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=DRG neuron;
RX   MEDLINE=22684492; PubMed=12771384;
RA   Tosetti P., Pathak N., Jacob M.H., Dunlap K.;
RT   "RGS3 mediates a calcium-dependent termination of G protein signaling
RT   in sensory neurons.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7337-7342(2003).
DR   EMBL; AY124776; AAM94022.1; -.
FT   NON_TER   1   1
SQ   SEQUENCE   441 AA; 48612 MW; F5BB2CC9C25EB70F CRC64;

Query Match   32.7%; Score 397; DB 13; Length 441;
Best Local Similarity   45.4%; Pred. No. 2.6e-23;
Matches   79; Conservative 38; Mismatches 55; Indels   2; Gaps   1;

Qy   31 EETSIEAKIR--AKEKRNRLSLLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFD 88
              ||||:|||| :||:::||:||:|:| |:||||:||::|||
Db   259 EETSRKKRSRSLAKDMKNRLGIFRRRNESPGANPSSKLDKVLKSLKPAPEEALKWGESLE 318

Qy   89 KLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAP 148
              :||:|| | :|::|: ||||||||||||||||||||||||||||||||||||||||||
Db   319 KLLLHKYGLAAFRAFLRTEFSEENLEFWLACEBYKKIKSQSKMVSKAKKIFAEYIAIQSC 378

Qy   149 KEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVQLMEHDSYKRFLKSETYLHLI 202
              ||||:||::|: :||: :: :|||||:||||||||||||||||||: ||:|:
Db   379 KEVNLDSYTREHTKENLQNITRGCFDLAQKRIYGLMEKDSYPRFLRSDLYLDII 432

RESULT 12
Q7SYH9
ID   Q7SYH9   PRELIMINARY;   PRT;   799 AA.
AC   Q7SYH9;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Regulator of G protein signaling 3 RGS3L isoform.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=DRG neuron;
RX   MEDLINE=22684492; PubMed=12771384;
RA   Tosetti P., Pathak N., Jacob M.H., Dunlap K.;
RT   "RGS3 mediates a calcium-dependent termination of G protein signaling
RT   in sensory neurons.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7337-7342(2003).
DR   EMBL; AY124777; AAM94023.1; -.
SQ   SEQUENCE   799 AA; 89511 MW; D0A818CF0EB822B2A CRC64;

Query Match   32.7%; Score 397; DB 13; Length 799;
Best Local Similarity   45.4%; Pred. No. 5.3e-23;
Matches   79; Conservative 38; Mismatches 55; Indels   2; Gaps   1;

Qy   31 EETSIEAKIR--AKEKRNRLSLLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFD 88
              ||||:|||| :||:::||:||:|:| |:||||:||::|||
Db   617 EETSRKKRSRSLAKDMKNRLGIFRRRNESPGANPSSKLDKVLKSLKPAPEEALKWGESLE 676

Qy   89 KLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAP 148
              :||:|| | :|::|: ||||||||||||||||||||||||||||||||||||||||||
Db   677 KLLLHKYGLAAFRAFLRTEFSEENLEFWLACEBYKKIKSQSKMVSKAKKIFAEYIAIQSC 736

Qy   149 KEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVQLMEHDSYKRFLKSETYLHLI 202
              ||||:||::|: :||: :: :|||||:||||||||||||||||||: ||:|:
Db   737 KEVNLDSYTREHTKENLQNITRGCFDLAQKRIYGLMEKDSYPRFLRSDLYLDII 790

RESULT 13
Q925G9
ID   Q925G9   PRELIMINARY;   PRT;   930 AA.
AC   Q925G9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   PDZ-RGS3 protein.
GN   RGS3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21197945; PubMed=11301003;
RA   Lu Q., Sun E.E., Klein R.S., Flanagan J.G.;
RT   "Ephrin-b reverse signaling is mediated by a novel PDZ-RGS protein and
RT   selectively inhibits G protein-coupled chemoattraction.";
RL   Cell 105:69-79(2001).
CC   -!- SIMILARITY: CONTAINS 1 RGS DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
DR   EMBL; AF350047; AAK38878.1; -.
DR   MGD; MGI:1354734; RGS3.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   InterPro; IPR001478; PDZ.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00595; PDZ; 1.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PD01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00228; PDZ; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50106; PDZ; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE   930 AA; 102511 MW; FB0ECEA164D2F6A1 CRC64;

Query Match   32.6%; Score 396; DB 11; Length 930;
Best Local Similarity   44.2%; Pred. No. 7.6e-23;
Matches   80; Conservative 38; Mismatches 61; Indels   2; Gaps   1;

Qy   31 EETSIEAKIR--AKEKRNRLSLLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFD 88
              :|||:|||| :||:::||:||:|:| |:||||:||::|||
Db   748 DEASRKRKSKNIAKDMKNKLAIFRRRNESPGAQPASKTDKTTKSFKPTSEEALKWSESLE 807

Qy   89 KLLSHRDGVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAP 148
              :||:|| | :|::|: ||||||||||||||||||||||||||: :||
Db   808 KLLLHKYGLEVFQAFLRTEFSEENLEFWLACEDFKKVKSQSKVAAKAKKIFAEFIAIQAC 867

Qy   149 KEVNIDFHTKEVIAKSIAQPTLHSFDTAQSRVQLMEHDSYKRFLKSETYLHLIEGRPQR 208
              ||||:||::|: :||: :: :|||||:|||||||||||||||||||||::||:
Db   868 KEVNLDSYTREHTKENLQSITRGCFDLAQKRIFGLMEKDSYPRFLRSDLYLDLINQKRMS 927

Qy   209 P 209
Db   928 P 928

RESULT 14
Q920Q9
ID   Q920Q9   PRELIMINARY;   PRT;   967 AA.
AC   Q920Q9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)

```
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   SRB-RGS.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21479395; PubMed=11595167;
RA   Ikeda M., Hirokawa M., Satani N., Kinoshita T., Watanabe Y., Inoue H.,
RA   Tone S., Ishikawa T., Minatogawa Y.;
RT   "Molecular cloning and characterization of a steroid receptor-binding
RT   regulator of G-protein signaling protein cDNA.";
RL   Gene 273:207-214 (2001).
CC   -!- SIMILARITY: CONTAINS 1 RGS DOMAIN
CC   -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
DR   EMBL; AB055153; BAB63460.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   InterPro; IPR001478; PDZ.
DR   InterPro; IPR000342; Regl_Gprotein.
DR   Pfam; PF00595; PDZ; 1.
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00228; PDZ; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50106; PDZ; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE   967 AA;  105380 MW;  5C558885DBA63A5B CRC64;

Query Match             32.2%;  Score 391;  DB 11;  Length 967;
Best Local Similarity   44.3%;  Pred. No. 2e-22;
Matches   77;  Conservative   37;  Mismatches   60;  Indels    0;  Gaps    0;

Qy   36 EAKIRAKEKRNRLSLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRD 95
           |: ||:  |||::|   :|:|| :  :||:|| :  ::|| ::| |||:|:
Db  792 KSKNIAKDMKNKLAIFRRRNESPGAQPAGKADKTTKSFKFPTSEEALKWSESLEKLLLHKY 851

Qy   96 GVDAFTRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDF 155
           :| |||||||||||||||          :: ||||||||||||||| | |||||
Db  852 GLEVFQAFLRTEFSEENLEFWLACEDFKKVKSQSKMAAKAKKIFAEFIAIQACKEVNLDS 911

Qy  156 HTKEVIAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEGRPQRP 209
           |:| :::  :|| |||:||  :| ||:::||:| |:|:|||||||:| ||
Db  912 YTREHTKENLQSITRGCFDLAQKRIFGLMEKDSYPRFLRSDLYLDLINQKKMSP 965


RESULT 15
Q8NFN6
ID   Q8NFN6      PRELIMINARY;      PRT;   319 AA.
AC   Q8NFN6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   RGS3 isoform RGS3S.
GN   RGS3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22032988; PubMed=12036301;
RA   Kehrl J.H., Srikumar D., Harrison K., Wilson G.L., Shi C.S.;
RT   "Additional 5' Exons in the RGS3 Locus Generate Multiple mRNA
RT   Transcripts, One of Which Accounts for the Origin of Human PDZ-RGS3.";
RL   Genomics 79:860-868(2002).
CC   -!- SIMILARITY: CONTAINS 1 RGS DOMAIN.
DR   EMBL; AF490838; AAM33253.1; -.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   InterPro; IPR000342; Regl_Gprotein.
```

```
DR   Pfam; PF00615; RGS; 1.
DR   PRINTS; PR01301; RGSPROTEIN.
DR   ProDom; PD001580; Regl_Gprotein; 1.
DR   SMART; SM00315; RGS; 1.
DR   PROSITE; PS50132; RGS; 1.
SQ   SEQUENCE   319 AA;  36094 MW;  998632C69B1F0F3 CRC64;

Query Match             31.6%;  Score 384;  DB 4;  Length 319;
Best Local Similarity   42.5%;  Pred. No. 1.9e-22;
Matches   79;  Conservative   34;  Mismatches   65;  Indels    8;  Gaps    1;

Qy   41 AKEKRNRLSLLQRPDFHGETQASRSALLAKETRVSPEEAVKWAESFDKLLSHRDGVDAF 100
          ||: ||:  |:|:  :| :|| : :  ::|| :: | |||:|:  :||| :|:
Db  141 AKDMKNKLGIFRRRNESPGAPPAGKADKMMKSFKFPTSEEALKWGESLEKLLVHKYGLAVF 200

Qy  101 TRFLKTEFSEENIEFWVACEDFKKCKEPQQIILKAKAIYEKFIQNDAPKEVNIDFHTKEV 160
          |||||||||||          :: ||||||||||||::|| :
Db  201 QAFLRTEFSEENLEFWLACEDFKKVKSQSKMASKAKKIFVEYIAIQACKEVNLDSYTREH 260

Qy  161 IAKSIAQPTLHSFDTAQSRVYQLMEHDSYKRFLKSETYLHLIEGRPQRPTNLRRRSRSFT 220
          ::  :|| |||:||  :| ||:::||:| |:|:|||||:|||      ::
Db  261 TKDNLQSVTRGCFDLAQKRIFGLMEKDSYPRFLRSDLYLDLI------NQKKMSPPLD 312

Qy  221 YNDFQD 226
          |: ||
Db  313 YKDDDD 318


Search completed: August 20, 2004, 16:47:00
Job time : 74.5 secs
```